Effective Field Theories for the LHC

by

Ian James Moult

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

Signature redacted

Author . . . .
Department of Physics
May 20, 2016

Signature redacted

Certified by . . . .
Iain W. Stewart
Professor of Physics
Thesis Supervisor

Signature redacted

Accepted by . . . .
Nergis Mavalvala
Associate Department Head of Physics

MITLibraries
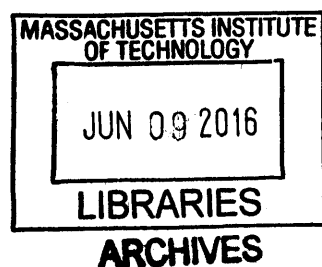
DISCLAIMER NOTICE

Due to the condition of the original material, there are unavoidable flaws in this reproduction. We have made every effort possible to provide you with the best copy available.

Thank you.

The images contained in this document are of the best quality available.

Effective Field Theories for the LHC

by

Ian James Moult

Submitted to the Department of Physics
on May 20, 2016, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

In this thesis I study applications of effective field theories to understand aspects of QCD jets and their substructure at the Large Hadron Collider. In particular, I introduce an observable, D_2, which can be used for distinguishing boosted $W/Z/H$ bosons from the QCD background using information about the radiation pattern within the jet, and perform a precision calculation of this observable. To simplify calculations in the soft collinear effective theory, I also develop a helicity operator basis, which facilitates matching calculations to fixed order computations performed using spinor-helicity techniques, and demonstrate its utility by computing an observable relevant for studying the properties of the newly discovered Higgs boson.

Thesis Supervisor: Iain W. Stewart
Title: Professor of Physics

Acknowledgments

Contents

List of Figures

16

17

19

20

21

22

24

List of Tables

27

Chapter 1

Introduction

The study of fundamental physics has progressed over the past century by probing the structure of particles at smaller and smaller scales, or equivalently by colliding them at higher and higher energies. In 2010, the Large Hadron Collider (LHC) became the highest energy collider, with a center of mass energy of 7 TeV (now 13 TeV), allowing for a probe of nature at unexplored scales. Already in its first few years of running, this has enabled the discovery of the Higgs boson.

Essential to maximizing the impact of the LHC program is a detailed theoretical understanding of the processes expected at the LHC within the Standard Model (SM) of particle physics. Precise theoretical predictions allow for the study of particles in the SM, as well as the detection of deviations from the SM, which would indicate the presence of new physics. Precision predictions at the LHC are made difficult by the fact that the LHC collides protons, which are not fundamental particles, but are made up of partons which interact strongly via Quantum Chromodynamics.

Quantum Chromodynamics is a non-abelian gauge theory. It is strongly coupled and confining at low energies, but becomes weakly coupled at high energies, allowing for a perturbative treatment of hard scattering processes. QCD has been extensively studied in hard scattering processes at e^+e^- colliders by measuring the pattern of radiation produced in collisions using observables called event shapes. Event shapes are theoretically clean, and

29
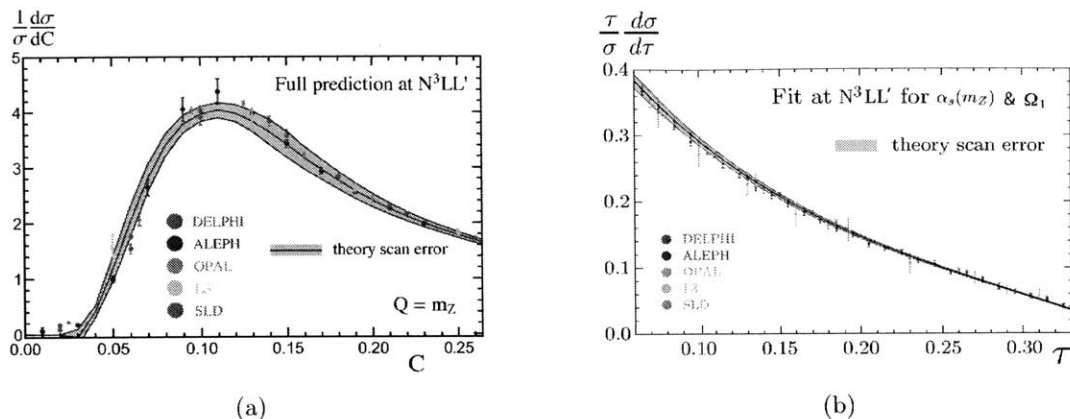
Figure 1-1: Precision theoretical predictions for the C-parameter (left) (figure from [1]) and Thrust (right) (figure from [2]) compared with data from e^+e^- colliders.

have been calculated to high precision. Two examples comparing theory predictions with experimental data are shown in Fig. 1-1 [2, 1].

The study of QCD at the extreme environments of the LHC is significantly more complicated, and requires a variety of new theoretical tools. This is due both to the strongly interacting nature of the colliding partons, as well as the level of sophistication of the measurements which can be performed on jets using the high granularity detectors at the LHC.

1.1 QCD at the LHC

The study of QCD plays an important role in nearly all measurements made at the LHC. Due to the fact that the colliding partons, namely quarks and gluons, are strongly interacting, QCD radiation, in the form of collimated, highly energetic jets is produced copiously. The study of the dynamics of these jets is extremely interesting for understanding the dynamics of QCD. Furthermore, electroweak particles such as the $W/Z/H$, are produced in association with QCD jets. Therefore the theoretical foundation for understanding any process at the LHC is dominated by QCD.

Consider a process with N final-state jets and L leptons, photons, or other nonstrongly
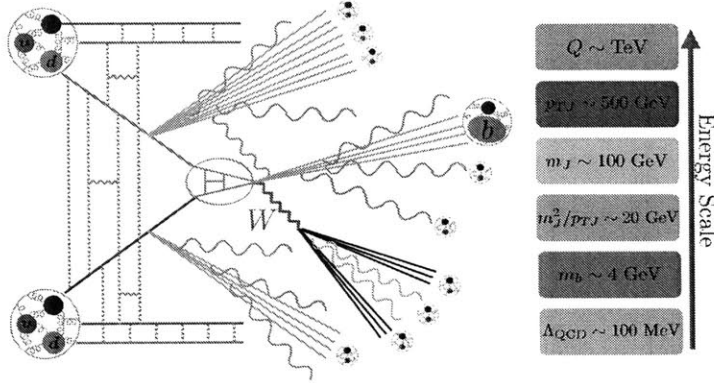
Figure 1-2: A schematic picture of the factorized structure of an LHC collision and the relevant scales in the problem. Jets with both resolved and unresolved substructure are illustrated.

interacting particles, with the underlying hard Born process

$$\kappa_a(q_a)\,\kappa_b(q_b) \to \kappa_1(q_1) \cdots \kappa_{N+L}(q_{N+L})\,, \tag{1.1}$$

where $\kappa_{a,b}$ denote the colliding partons, and κ_i denote the outgoing quarks, gluons, leptons, and other particles with momenta q_i. The incoming partons are essentially along the beam directions, $q_{a,b}^\mu = x_{a,b} P_{a,b}^\mu$, where $x_{a,b}$ are the momentum fractions and $P_{a,b}^\mu$ the (anti)proton momenta.

The active-parton exclusive jet cross section corresponding to Eq. (1.1) can be factorized for a variety of jet resolution variables.[1] The factorized expression for the exclusive jet cross section can be written schematically in the form

$$d\sigma = \int dx_a\, dx_b\, d\Phi_{N+L}(q_a + q_b; q_1, \ldots)\, M(\{q_i\}) \tag{1.2}$$
$$\times \sum_\kappa \text{tr}\left[\widehat{H}_\kappa(\{q_i\})\widehat{S}_\kappa\right] \otimes \left[B_{\kappa_a} B_{\kappa_b} \prod_J J_{\kappa_J}\right] + \cdots,$$

---

[1]Here active parton refers to initial-state quarks or gluons. Proofs of factorization with initial-state hadrons must also account for effects due to Glaubers [16], which may or may not cancel, and whose relevance depends on the observable in question [17, 18]. For a recent discussion in the context of SCET, see [19]

which is shown pictorially in Fig. 1-2. Here, $\mathrm{d}\Phi_{N+L}(q_a + q_b; q_1, \ldots)$ denotes the Lorentz-invariant phase space for the Born process in Eq. (1.1), and $M(\{q_i\})$ denotes the measurement made on the hard momenta of the jets (which in the factorization are approximated by the Born momenta q_i). The dependence on the underlying hard interaction is encoded in the hard function $\widehat{H}_\kappa(\{q_i\})$, where $\{q_i\} \equiv \{q_1, \ldots, q_{N+L}\}$, the sum over $\kappa \equiv \{\kappa_a, \kappa_b, \ldots \kappa_{N+L}\}$ is over all relevant partonic processes, and the trace is over color. Any dependence probing softer momenta, such as measuring jet masses or low p_Ts, as well as the choice of jet algorithm, will affect the precise form of the factorization, but not the hard function $\widehat{H}_\kappa$. This dependence enters through the definition of the soft function $\widehat{S}_\kappa$ (describing soft radiation), jet functions J_{κ_J} (describing energetic final-state radiation in the jets) and the beam functions B_i (describing energetic initial-state radiation along the beam direction). More precisely, the beam function is given by $B_i = \sum_{i'} \mathcal{I}_{ii'} \otimes f_{i'}$ with f_i the parton distributions of the incoming protons, and $\mathcal{I}_{ii'}$ a perturbatively calculable matching coefficient depending on the measurement definition [20]. The ellipses at the end of Eq. (1.2) denote power-suppressed corrections. All functions in the factorized cross section depend only on the physics at a single scale. This allows one to evaluate all functions at their own natural scale, and then evolve them to a common scale using their RGE. This procedure resums the large logarithms of scale ratios appearing in the cross section to all orders in perturbation theory. Furthermore, it enables the definition of universal non-perturbative quantities, such as the parton distribution functions, or fragmentation functions, which although they are not calculable from first principles, are universal, and thus predictivity is maintained. The description of even the simplest LHC observables therefore requires a detailed understanding of QCD, embodied in Eq. (1.2). The complexity becomes far greater as realistic experimental measurements are imposed on the final state.

One of the major advances at the LHC in the precision study of jets is the study of jet substructure, which will allow for an unprecedented study of strong dynamics. The extremely fine granularity of the LHC detectors allows for the substructure within a jet to be accurately measured, as opposed to simply its global properties. Experimentally, this is
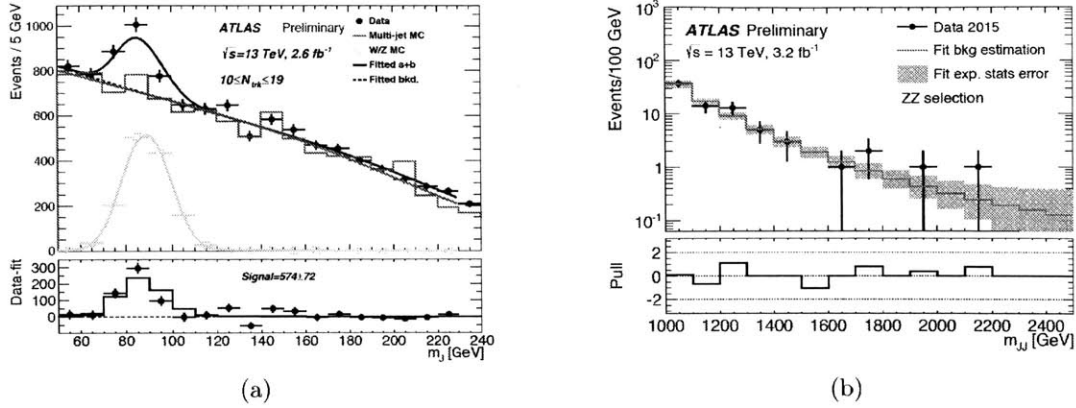
Figure 1-3: A plot of the invariant mass spectrum of jets produced at the LHC, showing the contribution from hadronically decaying W/Z bosons (left) and the dijet invariant mass spectrum used to search for resonances decaying to highly boosted W/Z bosons (right) (figures from [3]).

extremely useful, as the radiation pattern within a jet provides clues to the nature of the particle that created the jet. For example, a hadronically decaying $W/Z/H$ particle will decay primarily into a jet consisting of two hard prongs, while an average QCD jet produced by a radiating quark or gluon will not have such a structure. This is shown schematically in Fig. 1-2, where a hadronically decaying W jet is embedded into a complex LHC collision involving a number of other QCD jets. Similarly, jet substructure measurements can be used to determine the charge of a jet, to determine whether a jet originated from a quark or a gluon, or to identify jets coming from decaying top quarks. Jet substructure techniques have been used extensively in a variety of searches at the LHC. For example, they have been used to search for new heavy resonances decaying to boosted W/Z bosons, which then decay hadronically. The boosted W/Z bosons can be identified using jet substructure techniques, and the mass of the resonance can then be reconstructed. In Fig. 1-3 we show on the left a plot of the jet mass spectrum, showing a clear peak at the W/Z mass, and on the right, the dijet invariant mass spectrum obtained using jet substructure techniques [3]. New resonances would appear as bumps in this distribution. As a further example of the level of sophistication of the experimental techniques available for the study of jet substructure, in
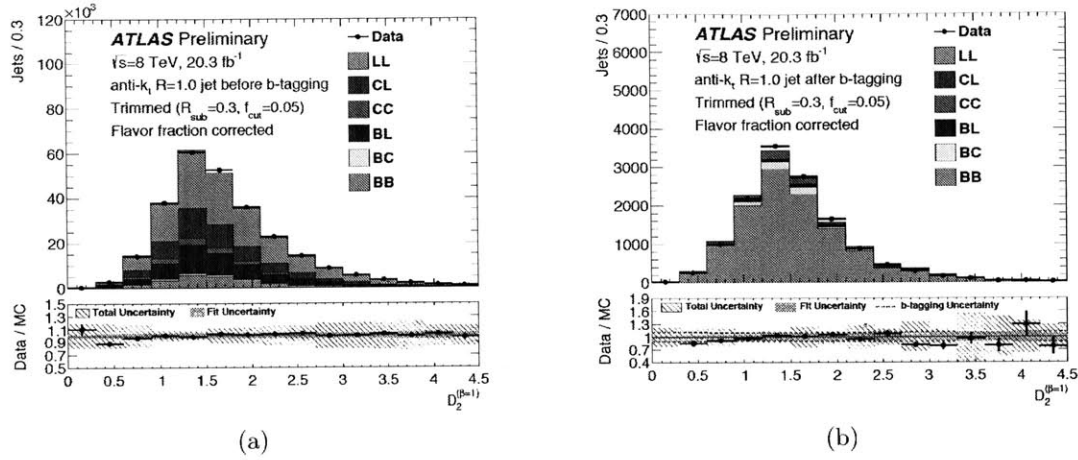
Figure 1-4: The distribution of radiation within a jet, as measured using D_2 before and after a subjet b-tagging is applied, showing the sophistication of modern jet substructure measurements (figures from [4]).

Fig. 1-4, we show a distribution of a jet shape observable D_2 [5], to be discussed in detail in this thesis, for gluon jets decaying to b quarks, $g \to b\bar{b}$ obtained using subjet b-tagging [4].

To be able to take advantage of these experimental measurements requires a detailed understanding of the substructure of QCD jets. One theoretically clean way that this can be achieved is to extend the ideas used in e^+e^- event shapes to jet shapes, which probe the radiation pattern within a particular jet, and which can be specifically designed to identify particular features of the jet. The calculation of jet shapes relevant for the LHC, for instance for boosted $W/Z/H$ production are still in their infancy. They are significantly more complicated than e^+e^- event shape calculations, as additional scales appear in the calculation associated with the substructure of the jet. To match the level of experimental sophistication in performing jet substructure measurements requires the development of new tools in QCD to accurately describe the substructure of QCD jets, and to extend the factorization formula of Eq. (1.2) to describe jet substructure observables.

34

1.2 Soft Collinear Effective Theory

Due to the complexity of QCD, effective field theories have played an important role, enabling calculations in specific limits of the theory. For example, effective field theories have been developed using power expansions in heavy quark masses, in the forward scattering limit, or more generally, in the presence of hierarchical kinematic scales. For studying QCD jets at the LHC, the appropriate effective field theory is the Soft Collinear Effective Field Theory (SCET).

SCET is an effective field theory of QCD describing the interactions of collinear and soft particles in the presence of a hard interaction, as relevant for jet production at the LHC [21, 22, 23, 24, 25]. In this section we briefly describe notation commonly used in SCET, give the Lagrangian of the theory, and show how field redefinitions can be used to factorize soft and collinear radiation, leading to an expression for a factorized cross section as in Eq. (1.2).

SCET describes collinear particles (which are characterized by a large momentum along a particular light-like direction), as well as soft particles. It is therefore convenient to use light-cone coordinates. For each jet direction we define two light-like reference vectors n_i^μ and $\bar{n}_i^\mu$ such that $n_i^2 = \bar{n}_i^2 = 0$ and $n_i \cdot \bar{n}_i = 2$. One typical choice for these quantities is

$$n_i^\mu = (1, \vec{n}_i), \qquad \bar{n}_i^\mu = (1, -\vec{n}_i), \tag{1.3}$$

where $\vec{n}_i$ is a unit three-vector. Given a choice for n_i^μ and $\bar{n}_i^\mu$, any four-momentum p can then be decomposed as

$$p^\mu = \bar{n}_i \cdot p \, \frac{n_i^\mu}{2} + n_i \cdot p \, \frac{\bar{n}_i^\mu}{2} + p_{n_i \perp}^\mu \ . \tag{1.4}$$

An "n_i-collinear" particle has momentum p close to the $\vec{n}_i$ direction, so that the components of p scale as $(n_i \cdot p, \bar{n}_i \cdot p, p_{n_i \perp}) \sim \bar{n}_i \cdot p \ (\lambda^2, 1, \lambda)$, where $\lambda \ll 1$ is a small parameter determined by the form of the measurement or kinematic restrictions. To ensure that n_i and n_j refer to distinct collinear directions, they have to be well separated, meaning

$$n_i \cdot n_j \gg \lambda^2 \qquad \text{for} \qquad i \neq j \ . \tag{1.5}$$

35

Two different reference vectors, n_i and n_i', with $n_i \cdot n_i' \sim \mathcal{O}(\lambda^2)$ both describe the same jet and corresponding collinear physics. Thus, each collinear sector can be labelled by any member of a set of equivalent vectors, $\{n_i\}$. This freedom is manifest as a symmetry of the effective theory known as reparametrization invariance (RPI) [26, 27]. Specifically, the three classes of RPI transformations are

$$
\begin{array}{lll}
\text{RPI-I} & \text{RPI-II} & \text{RPI-III} \\[1em]
n_{i\mu} \to n_{i\mu} + \Delta_\mu^\perp & n_{i\mu} \to n_{i\mu} & n_{i\mu} \to e^\alpha n_{i\mu} \\[1em]
\bar{n}_{i\mu} \to \bar{n}_{i\mu} & \bar{n}_{i\mu} \to \bar{n}_{i\mu} + \epsilon_\mu^\perp & \bar{n}_{i\mu} \to e^{-\alpha} \bar{n}_{i\mu} \,.
\end{array}
\tag{1.6}
$$

Here, we have $\Delta^\perp \sim \lambda$, $\epsilon^\perp \sim \lambda^0$, and $\alpha \sim \lambda^0$. The parameters $\Delta^\perp$ and $\epsilon^\perp$ are infinitesimal, and satisfy $n_i \cdot \Delta^\perp = \bar{n}_i \cdot \Delta^\perp = n_i \cdot \epsilon^\perp = \bar{n}_i \cdot \epsilon^\perp = 0$.

The effective theory is constructed by expanding momenta into label and residual components

$$
p^\mu = \tilde{p}^\mu + k^\mu = \bar{n}_i \cdot \tilde{p} \, \frac{n_i^\mu}{2} + \tilde{p}_{n_i\perp}^\mu + k^\mu \,.
\tag{1.7}
$$

Here, $\bar{n}_i \cdot \tilde{p} \sim Q$ and $\tilde{p}_{n_i\perp} \sim \lambda Q$ are the large label momentum components, where Q is the scale of the hard interaction, while $k \sim \lambda^2 Q$ is a small residual momentum. A multipole expansion is then performed to obtain fields with momenta of definite scaling, namely collinear quark and gluon fields for each collinear direction, as well as soft quark and gluon fields. Independent gauge symmetries are enforced for each set of fields.

The SCET fields for n_i-collinear quarks and gluons, $\xi_{n_i,\tilde{p}}(x)$ and $A_{n_i,\tilde{p}}(x)$, are labeled by their collinear direction n_i and their large momentum $\tilde{p}$. They are written in position space with respect to the residual momentum and in momentum space with respect to the large momentum components. Derivatives acting on the fields pick out the residual momentum dependence, $i\partial^\mu \sim k \sim \lambda^2 Q$. The large label momentum is obtained from the label momentum operator $\mathcal{P}_{n_i}^\mu$, e.g. $\mathcal{P}_{n_i}^\mu \xi_{n_i,\tilde{p}} = \tilde{p}^\mu \xi_{n_i,\tilde{p}}$. When acting on a product of fields, $\mathcal{P}_{n_i}$ returns the sum of the label momenta of all n_i-collinear fields. For convenience, we define $\overline{\mathcal{P}}_{n_i} = \bar{n} \cdot \mathcal{P}_{n_i}$, which picks out the large momentum component. Frequently, we will only keep

the label n_i denoting the collinear direction, while the momentum labels are summed over (subject to momentum conservation) and suppressed.

Soft degrees of freedom are described in the effective theory by separate quark and gluon fields. We will assume that we are working in the SCET$_\text{I}$ theory where these soft degrees of freedom are referred to as ultrasoft so as to distinguish them from the soft modes of SCET$_\text{II}$ [28]. In SCET$_\text{I}$, the ultrasoft modes do not carry label momenta, but have residual momentum dependence with $i\partial^\mu \sim \lambda^2 Q$. They are therefore described by fields $q_{us}(x)$ and $A_{us}(x)$ without label momenta. The ultrasoft degrees of freedom are able to exchange residual momenta between the jets in different collinear sectors. Particles that exchange large momentum of $\mathcal{O}(Q)$ between different jets are off-shell by $\mathcal{O}(n_i \cdot n_j Q^2)$, and are integrated out by matching QCD onto SCET. Before and after the hard interaction the jets described by the different collinear sectors evolve independently from each other, with only ultrasoft radiation between the jets.

SCET is formulated as an expansion in powers of λ, constructed so that manifest power counting is maintained at all stages of a calculation. As a consequence of the multipole expansion, all fields acquire a definite power counting [23], shown in Table 1.1. The SCET Lagrangian is also expanded as a power series in λ

$$\mathcal{L}_{\text{SCET}} = \mathcal{L}_{\text{hard}} + \mathcal{L}_{\text{dyn}} = \sum_{i \geq 0} \mathcal{L}_{\text{hard}}^{(i)} + \sum_{i \geq 0} \mathcal{L}^{(i)}, \qquad (1.8)$$

where (i) denotes objects at $\mathcal{O}(\lambda^i)$ in the power counting. The Lagrangians $\mathcal{L}_{\text{hard}}^{(i)}$ contain the hard scattering operators $O^{(i)}$, whose structure is determined by the matching process, as described in Sec. 5.4. The $\mathcal{L}^{(i)}$ describe the dynamics of ultrasoft and collinear modes in the effective theory, and their structure will be discussed shortly.

Factorization theorems used in jet physics are typically derived at leading power in λ. In this case, interactions involving hard processes in QCD are matched to a basis of leading power SCET hard scattering operators $O^{(0)}$, the dynamics in the effective theory are described by the leading power Lagrangian, $\mathcal{L}^{(0)}$, and the measurement function, which de-

Operator	$\mathcal{B}^{\mu}_{n_i\perp}$	χ_{n_i}	$\mathcal{P}^{\mu}_{\perp}$	q_{us}	D^{μ}_{us}
Power Counting	λ	λ	λ	λ^3	λ^2

Table 1.1: Power counting for building block operators in SCET$_{\rm I}$.

fines the action of the observable, is expanded to leading power. Higher power terms in the λ expansion, known as power corrections, arise from three sources: subleading power hard scattering operators $O^{(i)}$, subleading Lagrangian insertions, and subleading terms in the expansion of the measurement functions which act on soft and collinear radiation. The first two sources are independent of the form of the particular measurement, while the third depends on its precise definition.

In SCET, collinear operators are constructed out of products of fields and Wilson lines that are invariant under collinear gauge transformations [22, 23]. The smallest building blocks are collinearly gauge-invariant quark and gluon fields, defined as

$$\chi_{n_i,\omega}(x) = \left[\delta(\omega - \overline{\mathcal{P}}_{n_i})\, W^{\dagger}_{n_i}(x)\, \xi_{n_i}(x)\right], \tag{1.9}$$

$$\mathcal{B}^{\mu}_{n_i\perp,\omega}(x) = \frac{1}{g}\left[\delta(\omega + \overline{\mathcal{P}}_{n_i})\, W^{\dagger}_{n_i}(x)\, iD^{\mu}_{n_i\perp} W_{n_i}(x)\right].$$

With this definition of $\chi_{n_i,\omega}$, when expanded to a single quark, we have $\omega > 0$ for an incoming quark and $\omega < 0$ for an outgoing antiquark. For $\mathcal{B}_{n_i,\omega\perp}$, $\omega > 0$ ($\omega < 0$) corresponds to outgoing (incoming) gluons. In Eq. (5.15),

$$iD^{\mu}_{n_i\perp} = \mathcal{P}^{\mu}_{n_i\perp} + gA^{\mu}_{n_i\perp}, \tag{1.10}$$

is the collinear covariant derivative and

$$W_{n_i}(x) = \left[\sum_{\text{perms}} \exp\left(-\frac{g}{\overline{\mathcal{P}}_{n_i}}\,\bar{n}\cdot A_{n_i}(x)\right)\right], \tag{1.11}$$

is a Wilson line of n_i-collinear gluons in label momentum space. In general the structure of Wilson lines must be derived by a matching calculation from QCD. These Wilson lines sum

up arbitrary emissions of n_i-collinear gluons off of particles from other sectors, which due to the power expansion always appear in the $\bar{n}_i$ direction. The emissions summed in the Wilson lines are $\mathcal{O}(\lambda^0)$ in the power counting. The label operators in Eqs. (5.15) and (5.17) only act inside the square brackets. Since $W_{n_i}(x)$ is localized with respect to the residual position x, we can treat $\chi_{n_i,\omega}(x)$ and $\mathcal{B}_{n_i,\omega}^{\mu}(x)$ as local quark and gluon fields from the perspective of ultrasoft derivatives ∂^{μ} that act on x.

The complete set of collinear and ultrasoft building blocks for constructing hard scattering operators or subleading Lagrangians at any order in the power counting is given in Table 1.1. All other field and derivative combinations can be reduced to this set by the use of equations of motion and operator relations [29]. Since these building blocks carry vector or spinor Lorentz indices they must be contracted to form scalar operators, which also involves the use of objects like $\{n_i^{\mu}, \bar{n}_i^{\mu}, \gamma^{\mu}, g^{\mu\nu}, \epsilon^{\mu\nu\sigma\tau}\}$.

As shown in Table 1.1, both the collinear quark and collinear gluon building block fields scale as $\mathcal{O}(\lambda)$. For the majority of jet processes there is a single collinear field operator for each collinear sector at leading power. (For fully exclusive processes that directly produce hadrons there will be multiple building blocks from the same sector in the leading power operators since they form color singlets in each sector.) Also, since $\mathcal{P}_{\perp} \sim \lambda$, this operator will not typically be present at leading power (exceptions could occur, for example, in processes picking out P-wave quantum numbers). At subleading power, operators for all processes can involve multiple collinear fields in the same collinear sector, as well as $\mathcal{P}_{\perp}$ operator insertions. The power counting for an operator is obtained by adding up the powers for the building blocks it contains. To ensure consistency under renormalization group evolution the operator basis in SCET must be complete, namely all operators consistent with the symmetries of the problem must be included.

Dependence on the ultrasoft degrees of freedom enters the operators through the ultrasoft quark field q_{us}, and the ultrasoft covariant derivative D_{us}, defined as

$$ iD_{us}^{\mu} = i\partial^{\mu} + gA_{us}^{\mu}, \tag{1.12} $$

from which we can construct other operators including the ultrasoft gluon field strength. All operators in the theory must be invariant under ultrasoft gauge transformations. Collinear fields transform under ultrasoft gauge transformations as background fields of the appropriate representation. The power counting for these operators is shown in Table 1.1. Since they are suppressed relative to collinear fields, ultrasoft fields typically do not enter factorization theorems in jet physics at leading power. An example where ultrasoft fields enter at leading power is $B \to X_s \gamma$ in the photon endpoint region, which is described at leading power by a single collinear sector, and an ultrasoft quark field for the b quark.

The leading power SCET Lagrangian, $\mathcal{L}^{(0)}$, describing the interactions of soft and collinear particles in the effective theory can be written[2]

$$\mathcal{L}^{(0)} = \mathcal{L}^{(0)}_{n\xi} + \mathcal{L}^{(0)}_{ng} + \mathcal{L}^{(0)}_{us} , \qquad (1.13)$$

where

$$\mathcal{L}^{(0)}_{n\xi} = \bar{\xi}_n \Big(in \cdot D_{ns} + i\slashed{D}_{n\perp} W_n \frac{1}{\overline{\mathcal{P}}_n} W_n^\dagger i\slashed{D}_{n\perp} \Big) \frac{\slashed{\bar{n}}}{2} \xi_n , \qquad (1.14)$$

$$\mathcal{L}^{(0)}_{ng} = \frac{1}{2g^2} \text{tr}\big\{ ([iD^\mu_{ns}, iD^\nu_{ns}])^2 \big\} + \zeta \text{tr}\big\{ ([i\partial^\mu_{ns}, A_{n\mu}])^2 \big\} + 2\text{tr}\big\{ \bar{c}_n [i\partial^{ns}_\mu, [iD^\mu_{ns}, c_n]] \big\} ,$$

and the ultrasoft Lagrangian, $\mathcal{L}^{(0)}_{us}$, is simply the QCD Lagrangian. Here we have used a covariant gauge with gauge fixing parameter ζ for the collinear gluons. The various derivatives in Eq. (1.13) are defined as

$$iD^\mu_n = i\partial^\mu_n + gA^\mu_n , \qquad\qquad iD^\mu_{ns} = iD^\mu_n + \frac{\bar{n}^\mu}{2} gn \cdot A_{us} ,$$

$$i\partial^\mu_{ns} = i\partial^\mu_n + \frac{\bar{n}^\mu}{2} gn \cdot A_{us} , \qquad\qquad i\partial^\mu_n = \frac{\bar{n}^\mu}{2} n \cdot \partial + \frac{n^\mu}{2} \overline{\mathcal{P}} + \mathcal{P}^\mu_\perp , \qquad (1.15)$$

and include ultrasoft-collinear interactions.

A particularly convenient aspect of SCET is that the factorization of soft and collinear

---

[2]Here we assume that other possible modes, for example Glauber modes, do not contribute to the observable, and cancel out of the cross section.

40

degrees of freedom can be performed at the Lagrangian level using the BPS field redefinition. This decouples leading power interactions between soft and collinear particles, and moves them into Wilson lines appearing in the hard scattering operators. The BPS field redefinitions are given by

$$\mathcal{B}_{n\perp}^{a\mu} \to \mathcal{Y}_n^{ab} \mathcal{B}_{n\perp}^{b\mu}, \qquad \chi_n^\alpha \to Y_n^{\alpha\beta} \chi_n^\beta. \tag{1.16}$$

Additionally, ghost particles c_n, which are present after gauge fixing, transform under BPS field redefinitions according to

$$c_n^a \to \mathcal{Y}_n^{ab} c_n^b. \tag{1.17}$$

After performing the BPS field redefinition, we have

$$\mathcal{L}^{(0)\text{BPS}} = \mathcal{L}_{n\xi}^{(0)\text{BPS}} + \mathcal{L}_{ng}^{(0)\text{BPS}} + \mathcal{L}_{us}^{(0)}, \tag{1.18}$$

where the ultrasoft Lagrangian is unchanged. The collinear quark Lagrangian is given by

$$\mathcal{L}_{n\xi}^{(0)\text{BPS}} = \bar{\chi}_n \left(in \cdot \mathcal{D}_n + i\slashed{\mathcal{P}}_{n\perp} \frac{1}{\overline{\mathcal{P}}_n} i\slashed{\mathcal{P}}_{n\perp} \right) \frac{\slashed{\bar{n}}}{2} \chi_n, \tag{1.19}$$

where

$$i\mathcal{D}_n^\mu = W_n^\dagger i D_n^\mu W_n, \tag{1.20}$$

and the collinear gluon Lagrangian is given by

$$\mathcal{L}_{ng}^{(0)\text{BPS}} = \frac{1}{2g^2} \text{tr}\left\{ ([iD_n^\mu, iD_n^\nu])^2 \right\} + \zeta \text{tr}\left\{ ([i\partial_n^\mu, A_{n\mu}])^2 \right\} + 2\text{tr}\left\{ \bar{c}_n [i\partial_\mu^n, [iD_n^\mu, c_n]] \right\}, \tag{1.21}$$

explicitly showing that ultrasoft and collinear particles have been decoupled to leading power. After BPS field redefinition, since the leading power Lagrangian, the cross section can be written as a product of matrix elements, each involving only soft or collinear fields, as in

Eq. (1.2).

It is important to note that the BPS field redefinition does not, however, decouple ultra-soft and collinear interactions in the subleading Lagrangians. However, these contributions are power suppressed, and therefore one only needs to consider a fixed number of insertions of subleading Lagrangians, so that factorization is still manifest.

As presented here, SCET is appropriate for treating problems with a collinear scale, a soft scale, and a hard scale. In the context of jet substructure, where additional observables are measured on the jet, other scales may be present, and additional modes must be added to the effective field theory to treat such cases. This situation will be discussed in detail in Chap. 4. In such many scale problems, the power of the effective theory, which enables one to separate a problem into separate calculations at each scale, becomes even greater. The extension of SCET to treat various kinematic hierarchies is currently an active area of research.

Chapter 2

Overview

In this section I provide a brief outline of the material presented in this thesis, and discuss the significance of the work in the context of other work performed during my thesis. Several key results and figures from different sections throughout this thesis are reproduced here. This section also contains comments on material related to the contents of the papers that appeared after the publication of the papers, for example measurements at the LHC, or related work by other groups.

The main focus of the work presented in this thesis is divided into two separate categories: the development of effective field theories for jet substructure, and the use of effective field theories for the calculation of precision cross sections involving jets at the LHC. Each of these will be discussed in turn.

2.1 Effective Field Theories for Jet Substructure

As discussed in Chap. 1, one of the major advances in the study of jets at the LHC is the ability to study in detail the substructure of a jet. This necessitates new calculations in QCD. In this section I discuss my work in the area of jet substructure and put into context the work presented in Chap. 3 and Chap. 4.

Jet substructure observables play an important role at the LHC, as was reviewed in
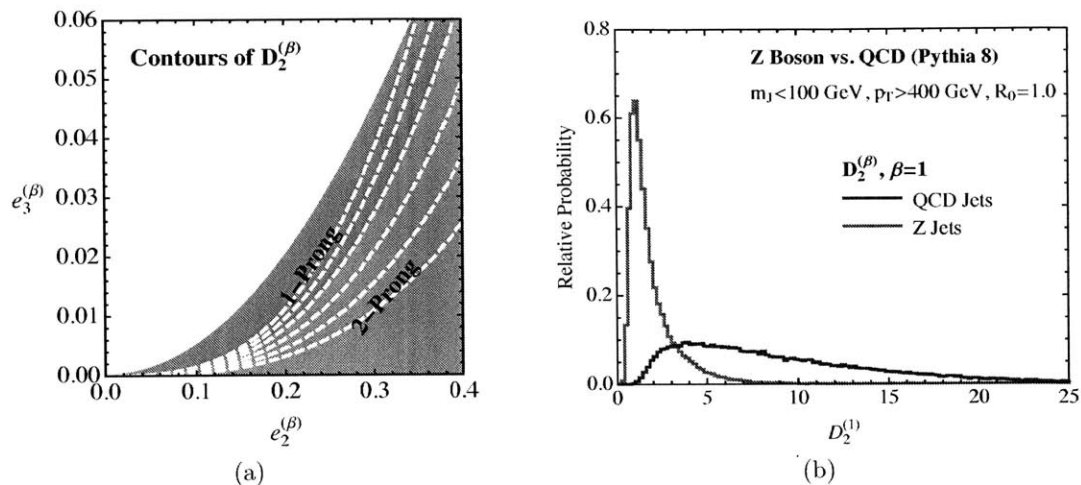
Figure 2-1: The phase space for $e_2^{(\beta)}$ and $e_3^{(\beta)}$ along with contours of constant $D_2^{(\beta)}$ separating the one- and two-prong regions of phase space (left). Monte Carlo distributions of the D_2 observable with $\beta = 1$ (right) (figures from [5]).

Chap. 1. These observables are very complicated to calculate from first principles, and therefore are most often studied using Monte Carlo programs. When designing jet substructure observables, a guess and check attitude is most often taken. It is therefore an interesting question to understand the extent that basic scaling arguments based on the fact that QCD is nearly a scale invariant theory can be used to understand the parametric behavior of jet substructure observables, and to design new jet substructure observables. After which, dedicated calculations, which are quite time intensive, can be performed. In "Power Counting to Better Jet Observables" [5], discussed in Chap. 3, we introduced the power counting approach for studying jet substructure observables formed from combinations of Infrared and Collinear Safe (IRC) safe observables. This approach allows one to make parametric predictions about the behavior of jet substructure observables, which are robust to tunings of Monte Carlo generators, or models of hadronization. We then used this approach to derive the two-prong jet substructure discriminant D_2, a ratio observable formed from the energy correlation functions $e_2^{(\beta)}$ and $e_3^{(\beta)}$ [30]. The observable D_2 is designed to discriminate boosted two-prong substructure from background QCD jets, by identifying the scaling of contours separating the $e_2^{(\beta)}$, $e_3^{(\beta)}$ phase space, as is shown in Fig. 2-1a. Its properties were

44
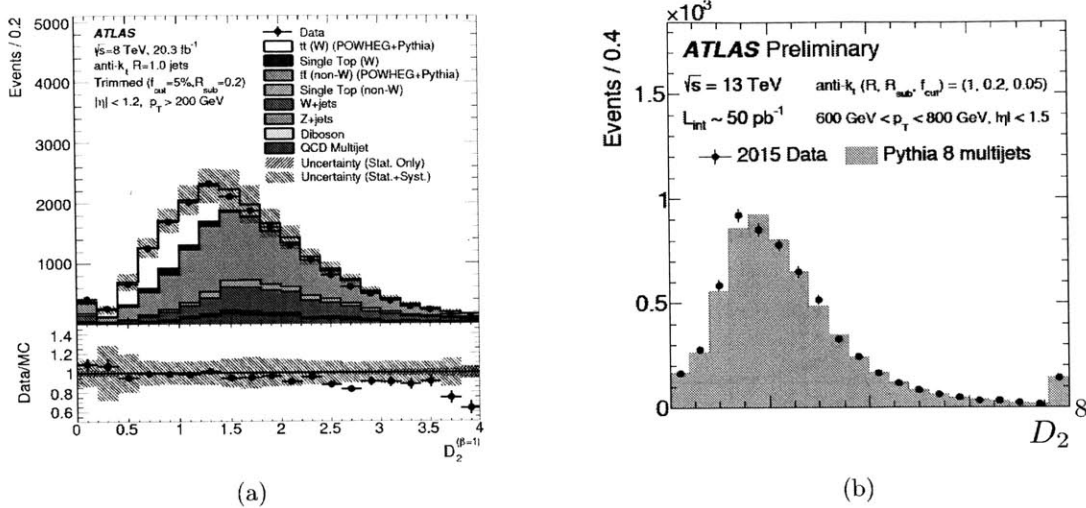
Figure 2-2: Measurements of the D_2 observable with angular exponent $\beta = 1$ by the ATLAS collaboration at the LHC at 8 TeV (left) (figure from [6]) and 13 TeV (right) (figure from [7]).

studied using power counting arguments and then compared with Monte Carlo simulations. An example Monte Carlo plot showing the D_2 observable as measured on QCD jets, and boosted Z jets, is shown in Fig. 2-1b.

The D_2 observable has been adopted by the ATLAS collaboration for tagging boosted two-prong substructure. Due to the importance of boosted two-prong tagging, for example for probing multi-TeV resonances, the D_2 observable was one of the first measurements at the 13 TeV LHC. In Fig. 2-2, measurements of the D_2 observable are shown at both 8 and 13 TeV as taken from [6, 7]. As mentioned in Chap. 1, quite interestingly, the D_2 distribution has also been measured on double b-tagged $g \to b\bar{b}$ jets [4]. This demonstrates some of the many advances made in the area of jet substructure made at the LHC in the past several years, including the ability to b-tag subjets, which can play an important role, for example, in boosted top tagging.

Top quarks play an important role in the Standard Model as the heaviest quark, with a mass of $m_t \sim 172$ GeV. Because of this large mass they play an important role in many models of BSM physics, where they are expected to couple strongly to new physics responsible for
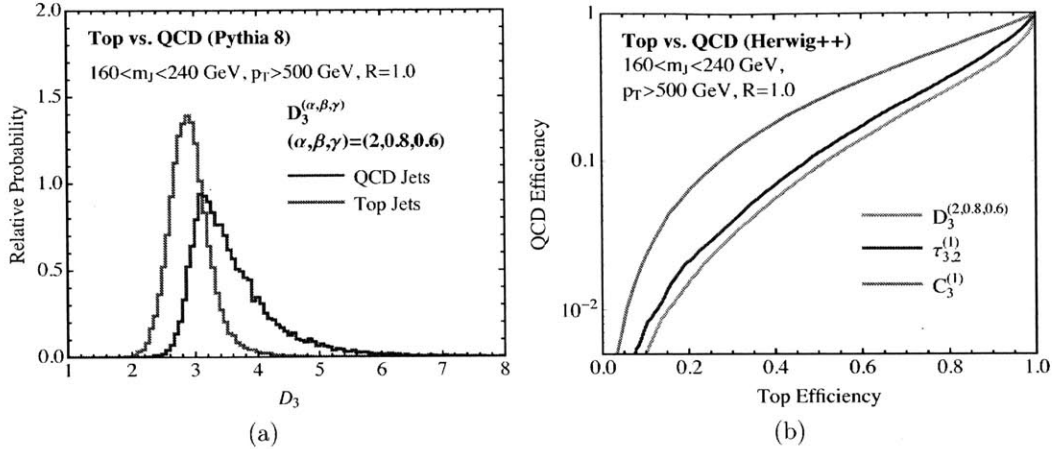
Figure 2-3: A distribution of $D_3^{(2,0.8,0.6)}$ as measured on top quark and QCD jets in Monte Carlo (left). Signal vs. Background efficiency curves comparing $C_3^{(1)}$, $D_3^{(2,0.8,0.6)}$, and $\tau_{3,2}^{(1)}$ (right) (figures from [8]).

mass generation. Such models of new physics often predict resonances in the TeV scale which could decay to highly boosted top quarks. Furthermore, at LHC energies, highly boosted top quarks are also produced copiously by Standard Model processes, and provide a probe of the dynamics of the Standard Model in the TeV regime. The ability to study boosted top quarks at the LHC relies on the ability to distinguish them from the QCD background, similar to the case of boosted $W/Z/H$ jets. A number of jet substructure variables have therefore been proposed for this purpose. In "Building a Better Boosted Top Tagger" [8], we extended the power counting arguments to identify the optimal boosted top tagger formed from the energy correlation functions. A plot of the proposed observable D_3 is shown in Fig. 2-3, and is compared with the standard N-Subjettiness observable $\tau_{3,2}$ [31], showing slightly improved performance. The D_3 observable is however, quite complicated, a reflection of the structure of the energy correlation function observables. It is currently being tested by the ATLAS collaboration. It would be interesting to identify a simpler boosted top quark observable formed from modified energy correlation functions.

The difficulty in performing analytic calculations of jet substructure observables, in particular those which are able to resolve an n-prong substructure within a jet, is that the
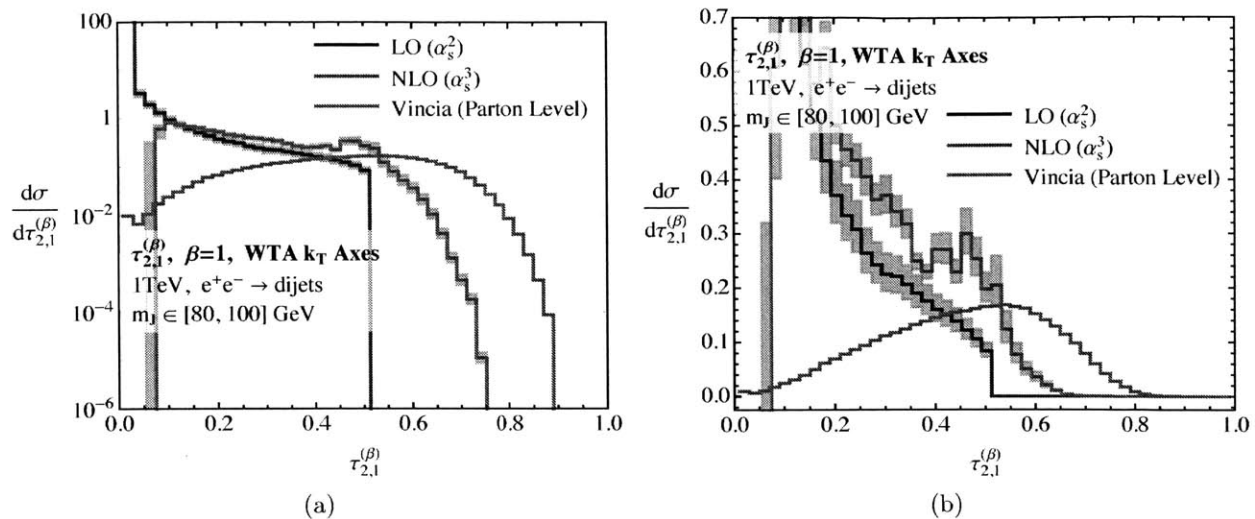
Figure 2-4: Distributions for $\tau_{2,1}$ as defined using WTA k_T axes with $\beta = 1$ in both fixed order perturbation theory and parton level Monte Carlo. The distributions are shown on a logarithmic scale in (a), and on a linear scale in (b). The behavior observed in Monte Carlo as $\tau_{2,1} \to 1$ is not well reproduced in fixed order perturbation theory (figures from [9]).

calculation must be able to describe the jet not only when it has n resolved subjets, but also when is has no resolved subjets, as well as for all intermediate options. Observables in QCD typically have singular behavior at edges of phase space regions, although it is also possible that observables exhibit singular behavior in the physical region. The location and structure of singular points for QCD observables plays an important role in their calculation, as naive perturbation theory breaks down in singular regions of phase space, and resummation is required. Observables typically also recieve large non-perturbative corrections in non-singular regions of phase space, which are difficult to calculate from first principles. It is therefore important to study the singular behavior of jet substructure observables.

In "The Singular Behavior of Jet Substructure Observables" [9], the singular behavior of two jet substructure observables D_2, and N-subjettiness [32, 31] were studied. The precise definition of these observables will be given in Chap. 3, but importantly the definition of the N-subjettiness observable requires a definition of the N-axes used to define the observable. A variety of different definitions have been proposed. One of the most powerful is the winner

take all axes definition [33]. We studied the N-subjettiness ratio observable for a variety of different axes definitions. This was performed using power counting arguments, as well as fixed order and parton shower Monte Carlo. In particular, it was shown that the singular behavior of the observable depends strongly on the definition of the N-subjettiness axes. In particular, depending on the definition of the axes, smooth distributions, singularities, or discontinuities can be observed in the distributions. Of particular phenomenological relevance, it was shown that with the winner take all axes definition, the N-jettiness ratio observable exhibits an infinite number of singularities in the physical region. This occurs because with a fixed number of partons, the winner take definition of N-subjettiness exhibits a maximal value, which can be surpassed by the addition of an extra parton. The first such shoulder is clearly seen in Fig. 2-4. This implies that the calculation in fixed order perturbation theory is extremely unstable. Furthermore, each of these singularities requires special treatment. It also implies the presence of large non-perturbative corrections throughout the entire distribution. This is in stark contrast to the behavior of the D_2 observable, which has a simple singular structure. In the presence of a mass cut, the D_2 observable isolates the singular region to $D_2 \to 0$, so that the effect of non-perturbative physics is isolated to this region of phase space, and is therefore controlled.

Having understood, using power counting techniques, the properties of a variety of popular jet substructure observables, as well as introducing observables with appealing theoretical properties, the next goal is the analytic calculation of jet substructure observables from first principles QCD. To achieve this goal, a number of new techniques needed to be introduced, namely the study of multi differential measurements, and the introduction of effective field theories describing jet substructure.

Many jet substructure observables, particularly those designed to tag jets with hierarchical scales, for example boosted W, Z, H jets, rely on the measurement of multiple observables on the same jet. Some common examples are the N-subjettiness ratio observables $\tau_{2,1}^{(\beta)}$ or the $D_2^{(\beta)}$. Theoretical calculations of these observables require calculations of multi-differential cross sections, that is cross sections differential in multiple jet measurements, for example,
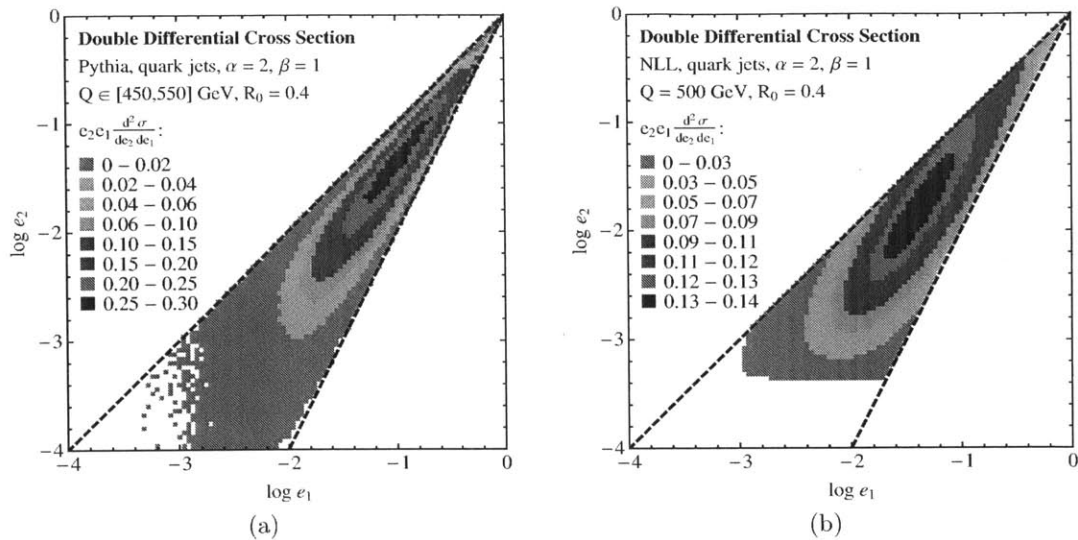
Figure 2-5: Plots of the double differential cross section defined from an analytic NLL interpolation measured on quark jets with one angularity fixed to be thrust ($\alpha = 2$) and scanning over the other angularity: $\beta = 1.5, 1, 0.5, 0.2$. The energy of the jets is $Q = 500$ GeV and the jet radius is $R_0 = 0.4$. The dashed lines on the plot correspond to the expected phase space boundary (figures from [10]).

the angularities with two different angular exponents.

In the paper "Toward Multi-Differential Cross Sections: Measuring Two Angularities on a Single Jet" [10], we initiated the study of the formal factorization and resummation properties of multi-differential cross-sections in QCD, by calculating the double differential cross section of two angularities $e_2^{(\alpha)}$ and $e_2^{(\beta)}$ measured on a single jet. A comparison of our calculation with Monte Carlo predictions is shown in Fig. 2-5. While these measurements do not resolve a two-prong substructure within the jet, as is required for the calculation of an observable such as D_2, they provide a simple example to understand how multi-differential cross sections can be studied in SCET. For example, in this paper, we noted that a new mode would need to be introduced into the effective theory, which is simultaneously soft and collinear. There has since been a number of works using soft-collinear modes in SCET, and studying multi-differential cross sections [34, 35, 36]. This work has also been used to study quark vs gluon tagging at the LHC [37].

49
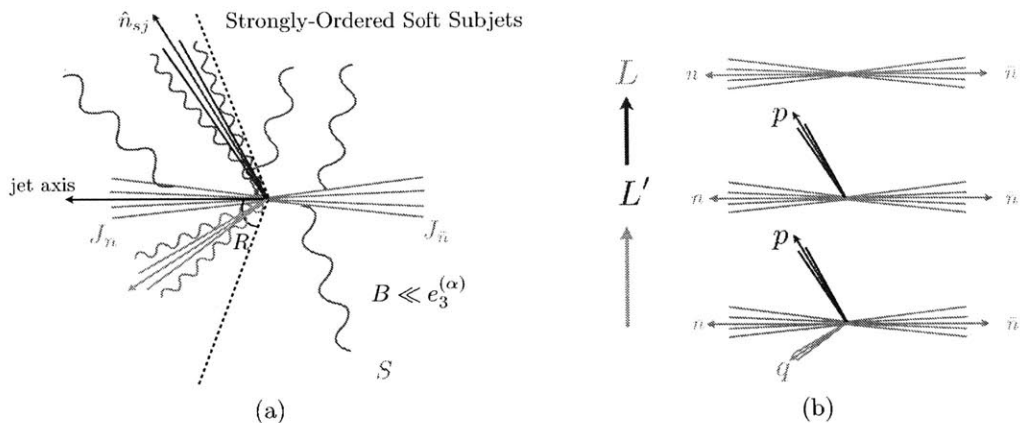
Figure 2-6: (a) Schematic depiction of the region of phase space defined by two strongly-ordered soft subjets. (b) Illustration of the resolved subjets as a function of the resolution scale, as implemented by the matching procedure in this region of phase space (figures from [11]).

In the paper, "Analytic Boosted Boson Discrimination" [12], discussed in Chap. 4 we performed the first rigorous factorization based analysis of a two prong jet substructure observable, and performed an explicit calculation for the case of the D_2 observable in $e^+e^- \to$ dijets, for both signal (boosted Z) and background (QCD) jets. This included a thorough analysis of the effective field theory descriptions of all phase space regions in the multi-differential phase space, a description of the non-perturbative effects on the D_2 distribution and an extensive comparison with a variety of Monte Carlo generators.

The structure of the D_2 observable derived from a power counting analysis played a central role in the calculability of the observable. Indeed, the form of the observable was chosen so that all phase space regions are parametrically separated, and so that contours of constant value of the D_2 observable are entirely contained in a single region of phase space. This allows the value of the differential cross section at each value of D_2 to be computed within a single effective field theory, greatly simplifying the calculation.

Our analysis indicated some discrepancies between the description of wide angle soft radiation in different Monte Carlo generators. The effective field theory description of this region of phase space was introduced in our paper "Non-Global Logarithms, Factorization,

50
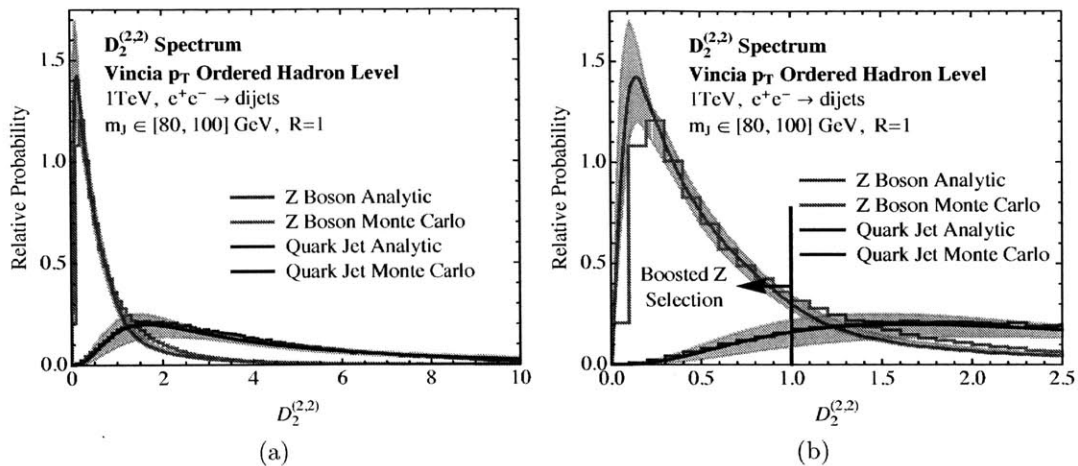
Figure 2-7: A comparison of signal and background $D_2^{(2,2)}$ distributions for the four different Monte Carlo generators and our analytic calculation, including hadronization. Here we show the complete distributions, including the long tail for the background distribution. Although we extend the factorization theorem beyond its naive region of applicability into the tail, excellent agreement with Monte Carlo is found (figures from [12]).

and the Soft Substructure of Jets" [11]. It allows for a description of the region of phase space where the jet consists of subjets with small energy, and wide angle with respect to the hard core of the jet, and allows a systematic separation of the different scales involved in this region of phase space. This is shown schematically in Fig. 2-6. This region of phase space also plays an important role in understanding Non-Global Logarithms [38]. We also proposed an event shape which specifically probes these non-global correlations, and performed an analytic calculation in e^+e^- [39].

In Fig. 2-7 a comparison of our analytic calculation including the effect of hadronization for the D_2 spectrum for boosted Z jets and QCD jets is shown and compared with Monte Carlo. Excellent agreement is observed. The uncertainties on the analytic prediction come from scale variations of the different functions appearing in the factorized formula for the D_2 cross sections. Details are given in Chap. 4. Having an explicit analytic prediction for both background and signal jets allowed us to make the first analytic ROC curves for a two-prong jet substructure observable, which are shown in Fig. 2-8. Again, overall good

51
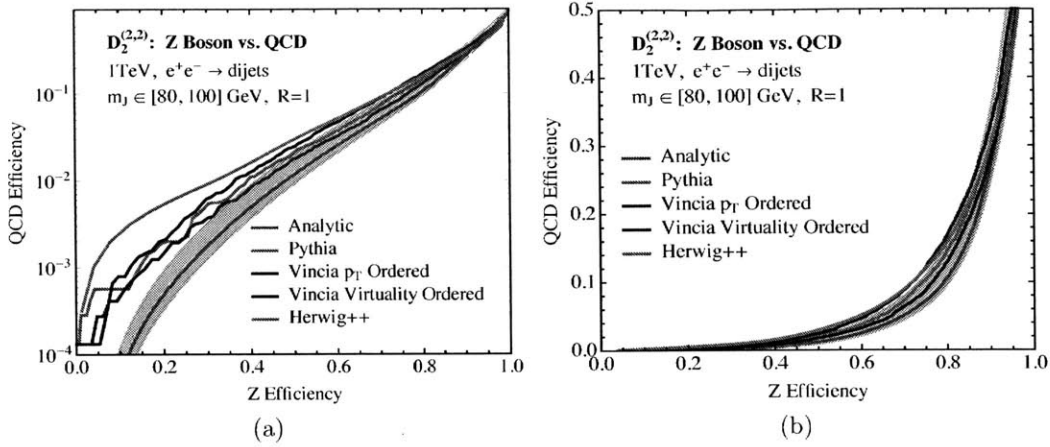
Figure 2-8: Signal vs. background efficiency curves for $D_2^{(2,2)}$ for the Monte Carlo samples as compared to our analytic prediction on a a) logarithmic scale plot and b) linear scale plot. The band of the analytic prediction is representative of the perturbative scale uncertainty. Good agreement between the analytic calculation and Monte Carlo is observed (figures from [12]).

agreement is observed, with the disagreement being driven by the signal distribution. It would be interesting to understand if higher order calculations would increase or reduce this discrepancy.

It would be of considerable practical and theoretical interest to extend this calculation to the LHC, where the D_2 observable is used extensively for boosted $W/Z/H$ tagging. There are two main obstructions to this calculation, one of which is experimental, and one of which is theoretical. On the experimental side, the D_2 observable is used in conjuction with a grooming procedure to remove the effects of pile-up. The effects of this grooming procedure would need to be incorporated in the theory calculation. On the theory side, the calculation in pp is considerably more complicated than in e^+e^- due to the color structure. For $e^+e^- \to$ dijets, even with an additional resolved jet, there are at most three jets, and a unique color structure, implying a color diagonal renormalization group evolution. For the LHC, even for the simplest process, such as $pp \to Z/W/H+$ jet, with an extra resolved subjet, there are four jets. This implies many partonic channels, each with non-diagonal

$H \to WW$

(a)

(b)
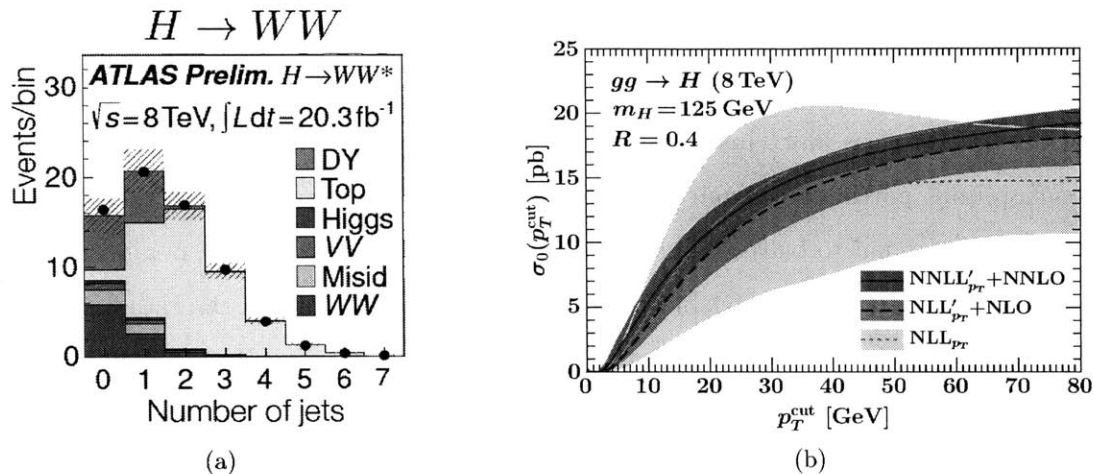
Figure 2-9: The distribution of signal and background processes in exclusive jet bins for the process $H \to WW$ (left). The exclusive zero jet spectrum for $gg \to H$ computed in resummed perturbation theory as a function of the p_T^{veto} used to define a jet (right)(figure from [13]).

color renormalization group evolution. It is interesting that grooming procedures, which are used to remove pile-up, also have the effect of removing wide angle color correlations. From the theory perspective, they therefore have the potential to simplify calculations. Because of these two considerations, it seems likely that the most interesting, and practical goal would be a calculation of a groomed D_2 at the LHC. This should be possible in the future using the groomer Soft Drop [40], which exhibits several nice theoretical properties [41, 42]. This would allow for first principles precision predictions for jet substructure variables of phenomenological importance at the LHC.

2.2 Precision Predictions for LHC Processes Involving Jets

One of the main goals of the LHC is the study of the newly discovered Higgs boson, as well as the W/Z bosons, and the top quark. Since the LHC is a proton-proton collider, these states are often produced in association with jets. The proper treatment of final states involving

53

jets is therefore essential for achieving precision predictions for the LHC. As an example, consider the process $H \rightarrow WW$ at the LHC. In Fig. 2-9, we show the distribution of signal and background process in exclusive jet bins. To separate the $H \rightarrow WW$ signal from the large background, particularly from top production, one can restrict to the exclusive zero jet bin where the signal to background ratio is highest. From a theoretical perspective, this restriction to the exclusive zero jet bin places significant constraints on the radiation in the final state, and introduces a new scale into the problem, namely p_T^{veto}, the scale above which a jet is defined as being a jet. The presence of this additional scale in the problem introduces large logarithms into the perturbative expansion, which need to be resummed to all orders to have precision predictions, as are required to measure properties of the Higgs boson. Such calculations have been performed in SCET for on-shell Higgs production, and the exclusive zero jet cross section is shown in Fig. 2-9 as a function of p_T^{veto} [13].

In performing such a calculation for a process involving jets at the LHC, many of the ingredients describing the dynamics of the jets are universal, and are described within the effective theory. The process dependence is carried by a matching coefficient describing the hard virtual fluctuations at the scale of the underlying hard process. The full QCD process is matched onto hard scattering operators in the effective field theory. To perform such a matching procedure, two ingredients are necessary. First, a complete basis of operators is required in the effective field theory, and second, the hard scattering amplitudes computed in fixed order QCD are required. For complicated processes, particularly those involving multiple jets, the construction of a complete operator basis in the effective field theory can be quite difficult, and furthermore, the fixed order amplitudes can be complicated, and often not in a form which can be easily interfaced with the SCET operators.

Many modern fixed order QCD calculations are performed using spinor helicity techniques for on-shell states of definite helicities. This has been shown to greatly simplify results, removing large gauge redundancies, and giving very compact expressions. The notion of helicity is also very natural in SCET, where external reference vectors identifying the directions of the jets are already present in the effective field theory. In the paper "Em-
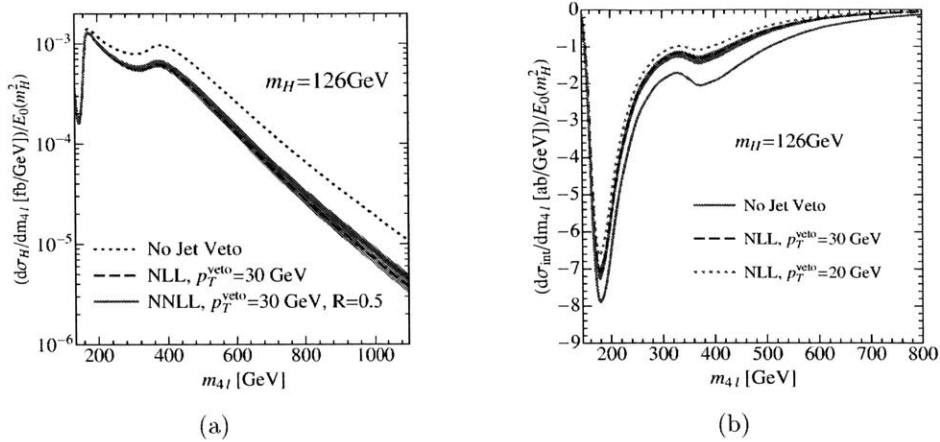
Figure 2-10: The off-shell Higgs cross section in the exclusive zero jet bin for the Higgs mediated process in (a) and the interference in (b) . Results are normalized by the jet veto suppression at the Higgs mass, such that the on-shell cross section is the same in all cases, allowing one to focus on the modification to the shape of the distribution. NLL and NNLL results are similar, with a small modification due to the finite jet radius, which is not present in the NLL calculation (figures from [14]).

ploying Helicity Amplitudes for Resummation" [43], discussed in Chap. 5, we introduced a convenient operator basis in SCET, which uses operators of definite helicity. These helicity operators make enumerating operator bases trivial, and the Wilson coefficients of the hard scattering operators are directly given by the (IR finite) pieces of color stripped helicity amplitudes. They allow state of the art predictions for the hard scattering process to be easily combined with an effective theory description of the dynamics of QCD jets. Hopefully this will enable more complex final states to be studied in SCET. Recently, we have also extended this formalism to subleading power in the SCET expansion [44].

To demonstrate the applicability of the helicity operators in SCET, in the paper "Jet Vetoes Interfering with $H \to WW$" [14], discussed in Chap. 6, I considered the resonance and continuum production of $pp \to l\nu \bar{l}\bar{\nu}$, which can proceed with through the Higgs boson, or through WW production. This process represents an interesting application of the helicity operators, as the amplitudes for $pp \to l\nu \bar{l}\bar{\nu}$ are complicated, and only available in spinor helicity form, and the matching to SCET can be greatly simplified. Furthermore, the process
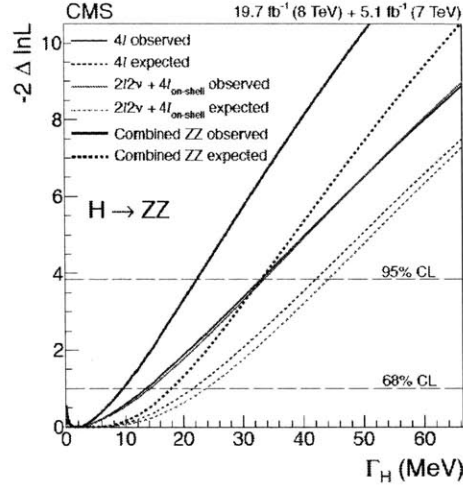
Figure 2-11: The constraints on the Higgs width as measured by CMS using the off-shell cross section for $H \to ZZ$ (figure from [15]).

is interesting as a probe of the properties of the Higgs. With the discovery of a Higgs like boson by both the ATLAS and CMS experiments at the LHC, an important goal of the LHC program will be the characterization of the resonance. Of particular importance is the width of the Higgs like boson. The width of the Standard Model boson is extremely narrow due to the fact that it is light, and couples proportional to mass. Other decay channels of the Higgs, could significantly modify the Higgs width. The experimental resolution on the direct lineshape is of the order of 1 GeV, while the predicted Standard Model value is ~ 4 MeV. Therefore, it is an important problem to come up with ways of measuring the Higgs width at the LHC. The width is observable at an e^+e^- machine through Higgstralung. It was noted that a (somewhat model dependent) way of bounding the Higgs width is to study its off shell production in the WW and ZZ channels and interference with continuum WW or ZZ production, due to the difference in scaling of the cross section with the width in the on-shell and off-shell region [45]. For the WW channel, the reduction of the background makes important use of jet vetoes. It is therefore an interesting question to understand how these jet vetoes effect the off-shell cross section.

56

In Chap. 6 we present a calculation of the exclusive zero jet cross section for both Higgs and continuum production, as well as their interference, in the far off-shell region. An interesting feature of this calculation, is that the size of the logarithm is a function of the invariant mass of the WW pair, and therefore changes throughout the invariant mass spectrum. This implies that the effect of the jet veto is not merely a rescaling of the cross section, but a reshaping of the distribution. This reshaping of the distribution is shown in Fig. 2-10. A precise calculation of this reshaping of the distribution is important in interpreting any possible new physics contributions to the spectrum. This measurement of the off-shell cross section has been performed at the LHC for the ZZ channel, and used to put a bound on the Higgs cross section. The result is shown in Fig. 2-11 [15].

Chapter 3

Power Counting to Better Jet Observables

Optimized jet substructure observables for identifying boosted topologies will play an essential role in maximizing the physics reach of the Large Hadron Collider. Ideally, the design of discriminating variables would be informed by analytic calculations in perturbative QCD. Unfortunately, explicit calculations are often not feasible due to the complexity of the observables used for discrimination, and so many validation studies rely heavily, and solely, on Monte Carlo. In this chapter we show how methods based on the parametric power counting of the dynamics of QCD, familiar from effective theory analyses, can be used to design, understand, and make robust predictions for the behavior of jet substructure variables. As a concrete example, we apply power counting for discriminating boosted Z bosons from massive QCD jets using observables formed from the n-point energy correlation functions. We show that power counting alone gives a definite prediction for the observable that optimally separates the background-rich from the signal-rich regions of phase space. Power counting can also be used to understand effects of phase space cuts and the effect of contamination from pile-up, which we discuss. As these arguments rely only on the parametric scaling of QCD, the predictions from power counting must be reproduced by any Monte Carlo, which we verify using PYTHIA 8 and HERWIG++. We also use the example of quark versus gluon

59

discrimination to demonstrate the limits of the power counting technique.

3.1   Introduction

Over the past several years there has been an explosion in the number of jet observables and techniques developed for discrimination and grooming [46, 47, 48]. Several of these are used by the ATLAS and CMS experiments, and their performance has been validated directly on data [49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70] and employed in new physics searches in highly boosted regimes [71, 72, 73, 74, 75, 76, 77, 78, 79]. Analyses using jets will become increasingly important at the higher energies and luminosities of Run 2 of the LHC.

While the proliferation of jet observables is exciting for the field, the vast majority of proposed observables and procedures have been analyzed exclusively in Monte Carlo simulation. Monte Carlos are vital for making predictions at the LHC, but should not be a substitute for an analytical understanding, where possible. Because Monte Carlos rely on tuning the description of non-perturbative physics to data, this can obscure what the robust perturbative QCD predictions are and hide direct insight into the dependence of the distributions on the parameters of the observable. This is especially confusing when different Monte Carlo programs produce different results.

Perturbative predictions of distributions have traditionally been constrained to only the simplest observables, such as the jet mass [80, 81, 82, 83, 84], but to high accuracy. Such high-order calculations are important for reducing the systematic theoretical uncertainties. More recently, resummation has been applied to some simple jet substructure variables [85, 86, 87, 37], and an understanding of some of the subtleties of resummation for ratio observables, as often used in jet substructure, has been developed [88, 10, 89]. Even the simplest calculations have suggested new, improved techniques, like the modified Mass Drop Tagger [86, 87], or uncovered unexpected structures in perturbative QCD, like Sudakov Safety [88, 40, 89]. For more complex observables, however, an analytic calculation may be essentially impossible, and we must rely on Monte Carlo simulations. Because of the

wide variety of jet observables, some of which can be calculated analytically and some that cannot, it is necessary to find an organizing principle that can be used to identify the robust predictions of QCD, without requiring a complete calculation to a given perturbative accuracy.

In this chapter we show how power counting methods can be used to design and understand the behavior of jet substructure variables. With minimal computational effort, power counting accurately captures the parametric predictions of perturbative QCD. The dynamics of a QCD jet are dominated by soft and collinear emissions and so by identifying the parametric scaling of soft and collinear contributions to a jet observable, we are able to make concrete and justified statements about the performance of jet substructure variables. Formal parametric scaling, or power counting, is widely used in the formalism of soft-collinear effective theory (SCET) [21, 22, 23, 24], an effective field theory of QCD in the soft and collinear limits. However, in this chapter, we will not rely on any results from SCET so as to make the discussion widely accessible. Similar techniques were employed in Ref. [90], but with the goal of determining which jet observables are calculable.

As a concrete application of the soft and collinear power counting method, we will focus on observables formed from the generalized n-point energy correlation functions $e_n^{(\beta)}$ [30], relevant for discriminating massive QCD jets from boosted, heavy objects. Measuring multiple energy correlation functions on a jet defines a multi-dimensional phase space populated by signal and background jets. By appropriately power counting the dominant regions of phase space, we are able to identify the signal- and background-rich regions and determine powerful observables for discrimination. In addition, from power counting arguments alone, we are able to predict the effect of pile-up contamination on the different regions of phase space. We apply power counting to the following:

- **Boosted Z Bosons vs. QCD** The two- and three-point energy correlation functions, $e_2^{(\beta)}$ and $e_3^{(\beta)}$, have been shown to be among the most powerful observables for identifying the hadronic decays of boosted Z bosons [30]. We discuss the phase space defined by $e_2^{(\beta)}$ and $e_3^{(\beta)}$, and determine which regions are populated by signal and background

61

jets. Using this understanding of the phase space, we propose a powerful discriminating variable to identify boosted two prong jets, given by

$$D_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3}.$$ (3.1)

This should be contrasted with the variable $C_2^{(\beta)} = e_3^{(\beta)}/(e_2^{(\beta)})^2$ originally proposed in Ref. [30]. We also show that power counting can be used to understand the impact of pile-up radiation on the different regions of phase space, and in turn to understand the susceptibility of signal and background distributions to pile-up.

- **Quarks vs. Gluons** Quark versus gluon jet discrimination is somewhat of a non-example for the application of power counting because there is nothing parametrically distinct between quark and gluon jets. However, this will illustrate why quark versus gluon discrimination is such a hard problem, and why different Monte Carlos can have wildly different predictions [37].

The outline of this chapter is as follows. In Sec. 3.2 we will precisely define what we mean by "collinear" and "soft" modes of QCD and introduce the observables used throughout this chapter. While we will mostly focus on the energy correlation functions, we will also discuss the N-subjettiness observables [32, 31] as a point of reference. In Sec. 3.3, we apply power counting to the study of Z versus QCD discrimination using the two- and three-point energy correlation functions $e_2^{(\beta)}$ and $e_3^{(\beta)}$. We argue that the single most powerful observable for discrimination is $e_3^{(\beta)}/(e_2^{(\beta)})^3$. Power counting is used to understand how the addition of pile-up radiation effects the distributions of this variable, and show that they are more robust to pile-up than for previously proposed variables formed from the energy correlation functions.[1] We verify that these predictions are borne out in Monte Carlo. In Sec. 3.4, we attempt to apply power counting to quark versus gluon jet discrimination. Naïvely, this should be the simplest case, however, power counting arguments are not applicable because all qualities

---

[1]The CMS study of Ref. [63] found that the observable $C_2^{(\beta)} \equiv e_3^{(\beta)}/(e_2^{(\beta)})^2$ suggested in Ref. [30] for boosted Z identification is very sensitive to pile-up contamination.

of quarks and gluons only differ by order-1 numbers. Finally, we conclude in Sec. 4.8 by re-emphasizing that power counting is a useful predictive tool for jet observables that are too complicated for direct analytic calculations, and suggest some problems to which it may prove fruitful.

3.2   Observable Basis and Dominant Physics of QCD

3.2.1   Observables

Throughout this chapter, our analyses will be focused around the (normalized) n-point energy correlation functions $e_n^{(\beta)}$.[2] The two-, and three-point energy correlation functions are defined as

$$
e_2^{(\beta)} = \frac{1}{p_{TJ}^2} \sum_{1 \leq i < j \leq n_J} p_{Ti} p_{Tj} R_{ij}^\beta \ ,
$$
$$
e_3^{(\beta)} = \frac{1}{p_{TJ}^3} \sum_{1 \leq i < j < k \leq n_J} p_{Ti} p_{Tj} p_{Tk} R_{ij}^\beta R_{ik}^\beta R_{jk}^\beta \ ,
$$

$$(3.3)$$

where p_{TJ} is the transverse momentum of the jet with respect to the beam, p_{Ti} is the transverse momentum of particle i, and n_J is the number of particles in the jet. The boost-invariant angle $R_{ij}^2 = (\phi_i - \phi_j)^2 + (y_i - y_j)^2$ is the Euclidean distance in the azimuth-rapidity plane and for infrared and collinear (IRC) safety, the angular exponent $\beta > 0$. In this chapter we will only study up through $e_3^{(\beta)}$, but higher-point energy correlation functions are defined as the natural generalization. We will often omit the explicit dependence on β, denoting the n-point energy correlation function simply as e_n.

---

[2]The notation $e_n^{(\beta)}$ differs from the original notation $\mathrm{ECF}(n,\beta)$ presented in Ref. [30] where the energy correlation functions were defined, but we hope that this notation used here is more compact. Specifically, the relationship is

$$
e_n^{(\beta)} = \frac{\mathrm{ECF}(n,\beta)}{(\mathrm{ECF}(1,\beta))^n} \ .
$$

$$(3.2)$$

The energy correlation functions have many nice properties that make them ideal candidates for defining a basis of jet observables. First, the energy correlation functions are defined such that $e_n^{(\beta)} \to 0$ in any of the soft or collinear limits of a configuration of n particles. Second, because all angles in the energy correlation functions are measured between pairs of particles, $e_n^{(\beta)}$ is insensitive to recoil or referred to as "recoil-free" [91, 92, 93, 30, 33]. This means that it is not sensitive to the angular displacement of the hardest particle (or jet core) from the jet momentum axis due to soft, wide angle radiation in the jet. The effects of recoil decrease the sensitivity of an observable to the structure of radiation about the hard core of the jet, making it less efficient for discrimination purposes.

Depending on the application, different energy correlation functions are useful as discriminating observables. As discussed in Ref. [30], the two-point energy correlation function is sensitive to radiation about a single hard core, and so is useful for quark versus gluon discrimination. Similarly, the three- and four-point energy correlation functions are useful for 2- or 3-prong jet identification, respectively, corresponding to boosted electroweak bosons $(W/Z/H)$ or hadronically decaying top quarks. By measuring appropriate energy correlation functions we define a phase space, populated by signal and background jets.

As a point of reference, we will also study the N-subjettiness observables and compare the structure of their phase space with that of the energy correlation functions. The (normalized) N-subjettiness observable $\tau_N^{(\beta)}$ is defined as

$$\tau_N^{(\beta)} = \frac{1}{p_{TJ}} \sum_{1 \leq i \leq n_J} p_{Ti} \min \left\{ R_{i1}^{\beta}, \ldots, R_{iN}^{\beta} \right\} . \tag{3.4}$$

The angle R_{iK} is measured between particle i and subjet axis K in the jet. Thus, N-subjettiness partitions a jet into N subjet regions and measures the p_T-weighted angular distribution with respect to the subjet axis of each particle. There are several different choices for how to define the subjet axes; here, we will define the subjet axes by the exclusive k_T jet algorithm [94] with the winner-take-all (WTA) recombination scheme [95, 33, 96]. In contrast to the traditional E-scheme recombination [97], which defines the (sub)jet axis to coincide with the net momentum direction, the WTA recombination scheme produces

(sub)jet axes that are recoil-free and nearly identical to the $\beta = 1$ minimized axes.[3] With this definition, the observables $e_2^{(\beta)}$ and $\tau_1^{(\beta)}$ are identical through NLL accuracy for all $\beta > 0$ [33].

Since N-subjettiness directly identifies N subjet directions in a jet, it is a powerful variable for N-prong jet discrimination. In particular, the N-subjettiness ratios

$$\tau_{2,1}^{(\beta)} \equiv \frac{\tau_2^{(\beta)}}{\tau_1^{(\beta)}} \quad \text{and} \quad \tau_{3,2}^{(\beta)} \equiv \frac{\tau_3^{(\beta)}}{\tau_2^{(\beta)}} \, ,$$

relevant for boosted $W/Z/H$ and top quark identification, respectively, are widely-used in jet studies at the ATLAS and CMS experiments. Numerical implementations of the energy correlation functions and N-subjettiness are available in the EnergyCorrelator and Nsubjettiness FASTJET CONTRIBs [100, 101].

3.2.2 Soft and Collinear Modes of QCD

At high energies, QCD is approximately a weakly-coupled conformal gauge theory and so jets are dominated by soft and collinear radiation. Because it is approximately conformal, there is no intrinsic energy or angular scale associated with this radiation. To introduce a scale, and so to determine the dominant soft and collinear emissions, we must break the conformal invariance by making a measurement on the jet. The scale of the soft and collinear emissions is set by the measured value of the observable.[4]

This observation can be exploited to make precise statements about the energy and angular structure of a jet, depending on the value of observables measured on that jet. This reasoning is often implicitly understood in the jet community and literature, and is formalized in SCET. Nevertheless, these precise power-counting arguments are not widely used outside of SCET, and so we hope that the applications in this chapter illustrate their

---

[3]The $\beta = 1$ minimized axes are also referred to as "broadening axes" [31, 33] as they correspond to axes that minimize the value of broadening [98, 99, 91].

[4]It is important to note that since QCD is not a conformal field theory, we can only use the power counting presented here to study the phase space defined by a set of IRC safe observables. If we considered IRC unsafe observables, then generically, we would need to power count contributions from non-perturbative physics such as hadronization.

effectiveness and relative simplicity.

We begin by defining a soft emission, s, as one for which

$$z_s \equiv \frac{p_{Ts}}{p_{TJ}} \ll 1 \; , \qquad R_{sj} \sim 1 \; , \qquad (3.5)$$

where j is any other particle in the jet and $R_{sj} \sim 1$ means that R_{sj} is not associated with any parametric scaling. Similarly, a collinear emission, c, is defined as having a p_T fraction

$$\frac{p_{Tc}}{p_{TJ}} \sim 1 \; , \qquad (3.6)$$

but with an angle to other particles which depends on whether they are also collinear or soft:

$$R_{cc} \ll 1 \; , \qquad R_{cs} \sim 1 \; . \qquad (3.7)$$

Here, R_{cc} is the angle between two collinear particles, while R_{cs} is the angle between a soft particle and a collinear particle. The precise scalings of R_{cc} and z_s will depend on the observable in question, as will be explained shortly. Soft emissions also implicitly include radiation that is simultaneously both soft and collinear.

To introduce these ideas concretely, we use the example of the two-point energy correlation function:

$$e_2^{(\beta)} = \frac{1}{p_{TJ}^2} \sum_{1 \leq i < j \leq n_J} p_{Ti} p_{Tj} R_{ij}^{\beta} \; . \qquad (3.8)$$

Consider performing a measurement of $e_2^{(\beta)}$ on a jet and further requiring $e_2^{(\beta)} \ll 1$. Because the energy flow in jets is in general collimated, this defines a non-trivial region of phase space, with a large fraction of jets satisfying this requirement. A large value of $e_2^{(\beta)}$ would mean that there is a hard, perturbative splitting in the jet which is suppressed by the small value of α_s. From the definition of $e_2^{(\beta)}$ in Eq. (3.8), we see that a measurement of $e_2^{(\beta)} \ll 1$ forces all particles in the jet to either have small p_{Ti} or small R_{ij}. In other words, the observable is dominated by soft and collinear emissions. The precise scaling of p_{Ti} and R_{ij}

is then determined by the measured value of $e_2^{(\beta)}$.[5]

There are three possible configurations that contribute to $e_2^{(\beta)} \ll 1$: soft-soft correlations, soft-collinear correlations, and collinear-collinear correlations. Therefore, $e_2^{(\beta)}$ can be expressed as

$$e_2^{(\beta)} \sim \frac{1}{p_{TJ}^2} \sum_s p_{Ts} p_{Ts} R_{ss}^\beta + \frac{1}{p_{TJ}^2} \sum_{s,c} p_{Ts} p_{Tc} R_{cs}^\beta + \frac{1}{p_{TJ}^2} \sum_c p_{Tc} p_{Tc} R_{cc}^\beta , \qquad (3.9)$$

where we have separated the contributions to $e_2^{(\beta)}$ into the three different correlations. To determine the dominant contributions to $e_2^{(\beta)}$, we will throw away those contributions that are parametrically smaller, according to our definitions of soft and collinear above. First, $p_{Ts} \ll p_{Tc}$, and so we can ignore the first term to leading power. Because $R_{cs} \sim 1$, we set $R_{cs} = 1$ in the second term, for the purpose of scaling. Also, note that $p_{Tc} \sim p_{TJ}$ and so we can replace the instances of p_{Tc} with p_{TJ} in the second and third terms. Making these replacements, we find

$$e_2^{(\beta)} \sim \sum_s z_s + \sum_c R_{cc}^\beta , \qquad (3.10)$$

where we have ignored any corrections arising at higher power in the soft and collinear emissions' energies and angles. We wish to emphasize with the explicit summation symbols that we have not restricted to a single soft or collinear emission, but consider an arbitrary number of emissions. Furthermore, we do not assume a strongly ordered limit, but instead explore the complete phase space arising from soft and collinear emissions, including regions where such ordering is explicitly broken.

Eq. (3.10) demonstrates the dominant structure of a jet on which we have measured $e_2^{(\beta)} \ll 1$. The contribution to $e_2^{(\beta)}$ from soft and collinear emissions do not mix to this accuracy; that is, they factorize from one another. Also, because there is no measurement

---

[5]In SCET, R_{cc} and z_s are often immediately assigned a related scaling. While this is true for this example, it is not in general true in the case of multiple measurements, and we wish to emphasize in this section how the measurement sets both scalings.

to distinguish the soft and collinear contributions to $e_2^{(\beta)}$, we then have that

$$e_2^{(\beta)} \sim z_s \sim R_{cc}^\beta\,. \qquad (3.11)$$

That is, the measured value of $e_2^{(\beta)}$ sets the p_T of the soft particles relative to the jet and the splitting angle of the collinear particles, and therefore defines the structure of the jet.

In Eq. (3.10), we have explicitly written a summation over the particles with soft and collinear scalings. To determine the scalings of the different contributions, it is clearly sufficient to consider the scaling of an individual term in each sum. In the remainder of this chapter we will drop the explicit summation for notational simplicity.

Scaling arguments similar to the power counting approach discussed here are often used in other approaches to QCD resummation to identify the relevant soft and collinear scales, and could also be used to analyze observables. For example, in the method of regions [102, 103], the regions of integration over QCD matrix elements which contribute dominantly to a given observable are determined, and an expansion about each of these regions is performed. These regions of integration, and the scaling of the momenta in these regions, basically correspond to the modes of the effective theory determined through the power counting approach.[6]

Similarly, in the CAESAR approach to resummation [93], implemented in an automated computer program, the first step of the program is the identification of the relevant soft and collinear scales. This is performed by expanding a given observable in the soft and collinear limits, and considering the region of integration for a single emission. This procedure is similar to that used in the case of $e_2^{(\beta)}$ just discussed, and would identify the same dominant contributions and scalings. Using the knowledge of the behavior of the QCD splitting functions, CAESAR then performs a resummed calculation of the observable. However, the CAESAR computer program is currently restricted to observables for which the relevant scales, and hence the logarithmic structure, is determined by a single emission, and further, to single differential distributions.

When considering observables relevant for jet substructure, one is interested in variables

---

[6]More precisely, only on-shell modes appear as degrees of freedom in the effective theory.

such as $e_n^{(\beta)}$, $n > 2$, whose behavior is not determined by the single emission phase space. For such observables, the single-emission analysis is not sufficient and the explicit analysis of QCD matrix elements to determine the dominant regions of integration which contribute becomes quite complicated. For these cases, we find the power counting approach of the effective field theory paradigm to be a particularly convenient organizing principle. Using the knowledge that on-shell soft and collinear modes dominate, a consistent power counting can be used to determine the relevant scalings of these modes in terms of the measured observables, which is reduced to a simple algebraic exercise. Although the evaluation of QCD matrix elements in these scaling limits is of course required for a complete calculation, it is not required to determine the power counting, and we will see that power counting alone will often be sufficient for constructing discriminating observables for jet substructure studies.

Throughout the rest of this chapter, we will employ these power-counting arguments to determine the dominant structure of jets on which multiple measurements have been made, for example $e_2^{(\beta)}$ and $e_3^{(\beta)}$. In this case, the phase space that results is much more complicated than the example of $e_2^{(\beta)}$ discussed above, but importantly, appropriate power counting of the contributions from soft and collinear emissions will organize the phase space into well-defined regions automatically.

3.3 Power Counting Boosted Z Boson vs. QCD Discrimination

As a detailed example of the usefulness of power counting, we consider the problem of discriminating hadronically-decaying, boosted Z bosons from massive QCD jets. Because Z boson decays have a 2-prong structure, we will measure the two- and three-point energy correlation functions, e_2 and e_3, on the jets, defining a two-dimensional phase space. We will find that there are two distinct regions of this phase space corresponding to jets with one or two hard prongs. QCD jets exist dominantly in the former region while boosted Z bosons

exist dominantly in the latter. Power counting these phase space regions will allow us to determine the boundaries of the regions and to define observables that separate the signal and background regions most efficiently.

Both because it is a non-trivial application, as well as still being tractable, we will present a detailed analysis of the phase space regions for boosted Z identification. This will require several pieces. First, we will study the full phase space of perturbative jets defined by e_2 and e_3 and identify signal and background regions via power counting. This will lead us to define a discriminating variable, $D_2^{(\beta)}$. Second, any realistic application of a boosted Z tagger includes a cut on the jet mass in the window around m_Z, and the effect of the mass cut on the discrimination power can also be understood by a power counting analysis of the phase space. Third, at the high luminosities of the LHC, contamination from pile-up is important and can substantially modify distributions for jet substructure variables. By appropriate power counting of the pile-up radiation, we can understand the effect of pile-up on the perturbative phase space and determine how susceptible the distributions of different discrimination variables are to pile-up contamination. As a reference, throughout this section we will contrast the energy correlation functions to the N-subjettiness observables $\tau_1^{(\beta)}$ and $\tau_2^{(\beta)}$ [32, 31]. A full effective theory analysis and analytic calculation of $D_2^{(\beta)}$ will be presented in Chap. 4.

3.3.1 Perturbative Radiation Phase Space

We begin by studying the (e_2, e_3) phase space arising from perturbative radiation from the jet. The measurement of e_2 and e_3 on a jet can resolve at most two hard subjets. The phase space for the variables e_2 and e_3 is therefore composed of jets which are unresolved by the measurement, dominantly from the QCD background, and shown schematically in Fig. 3-1a, and jets with a resolved 2-prong structure, as from boosted Z decays, shown schematically in Fig. 3-1b. We will find that the resolved and unresolved jets live in parametrically different regions of the phase space, and the boundary between the two regions can be understood from a power counting analysis.

70

(a)   (b)
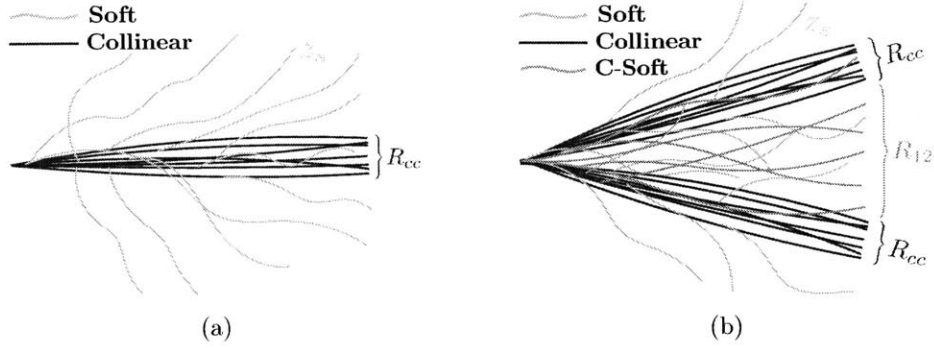
Figure 3-1: a) 1-prong jet, dominated by collinear (blue) and soft (green) radiation. The angular size of the collinear radiation is R_{cc} and the p_T fraction of the soft radiation is z_s. b) 2-prong jet resolved into two subjets, dominated by collinear (blue), soft (green), and collinear-soft (orange) radiation emitted from the dipole formed by the two subjets. The subjets are separated by an angle R_{12} and the p_T fraction of the collinear-soft radiation is z_{cs}.

modes	$e_2^{(\beta)}$	$e_3^{(\beta)}$
CCC	R_{cc}^{β}	$R_{cc}^{3\beta}$
CCS	$R_{cc}^{\beta} + z_s$	$z_s R_{cc}^{\beta}$
CSS	$z_s + z_s^2$	z_s^2
SSS	z_s^2	z_s^3

Table 3.1: Scaling of the contributions of 1-prong jets to $e_2^{(\beta)}$ and $e_3^{(\beta)}$ from the different possible configurations of soft (S) and collinear (C) radiation.

First, consider the case of the measurement of $e_2^{(\beta)}$ and $e_3^{(\beta)}$ on a jet with a single hard core of radiation, as in Fig. 3-1a, which is dominated by soft radiation with characteristic p_T fraction $z_s \ll 1$, and collinear radiation with a characteristic angular size $R_{cc} \ll 1$. All other scales are order-1 numbers that we will assume are equal to 1 without further discussion. With these assumptions, we are able to determine the scaling of the contributions to $e_2^{(\beta)}$ and $e_3^{(\beta)}$ from collections of soft and collinear particles. The scalings are given in Table 3.1 for contributions from three collinear particles (CCC), two collinear and one soft particle (CCS), one collinear and two soft particles (CSS), and three soft particles (SSS).[7]

---

[7]The contributions in Table 3.1 are from *any* subset of three particles in the jet. We do not single out an initial parton from which the others arise as in a showering picture.

Dropping those contributions that are manifestly power-suppressed, the two- and three-point energy correlation functions measured on 1-prong jets therefore scale like

$$e_2^{(\beta)} \sim R_{cc}^{\beta} + z_s \,, \tag{3.12}$$

$$e_3^{(\beta)} \sim R_{cc}^{3\beta} + z_s^2 + R_{cc}^{\beta} z_s \,. \tag{3.13}$$

To go further, we must determine the relative size of z_s and R_{cc}^{β}. There are two possibilities, depending on the region of phase space identified by the measurement: either z_s makes a dominant contribution to e_2, or its contribution is power suppressed with respect to R_{cc}^{β}. In the case that z_s contributes to e_2, this immediately implies that $e_3 \sim (e_2)^2$, regardless of the precise scaling of R_{cc}^{β}.[8] If instead z_s gives a subleading contribution compared to R_{cc}^{β} in e_2, then $e_3 \sim (e_2)^3$.[9] Therefore, from this simple analysis, we have shown that 1-prong jets populate the region of phase space defined by $(e_2)^3 \lesssim e_3 \lesssim (e_2)^2$. Fascinatingly, this also implies that the relative values of e_2 and e_3 provide a direct probe of the ordering of emissions inside the jet, so that assumptions about the measured values of e_2 and e_3 are observable proxies for the ordering of emissions. The scaling of R_{cc} and z_s on each boundary of the phase space can then easily be determined, but will not be important for our discussion.

This analysis shows that 1-prong jets fill out a non-trivial region in the (e_2, e_3) phase space, and of particular interest for the design of discriminating observables is the fact that this region of phase space has a lower boundary. This region is shown in blue in Fig. 3-2. To understand the region of phase space for $e_3 \ll (e_2)^3$ we must consider the case in which the measurement of e_2 and e_3 resolves two subjets within the jet.

The setup for the power counting of 2-prong jets is illustrated in Fig. 3-1b. We consider a jet with two subjets, each of which carry $\mathcal{O}(1)$ of the jet p_T and are separated by an angle $R_{12} \ll 1$. Each of the subjets has collinear emissions at a characteristic angle $R_{cc} \ll R_{12}$. Because $R_{12} \ll 1$, there is in general global soft radiation at large angles with respect to the

---

[8]Note that on the true upper boundary of the phase space, the assumption of strong ordering of emissions is broken.

[9]The existence of a consistent power counting does not guarantee a factorization theorem. Indeed, while a factorization theorem exists on the quadratic boundary, it does not exist at leading power on the cubic boundary.
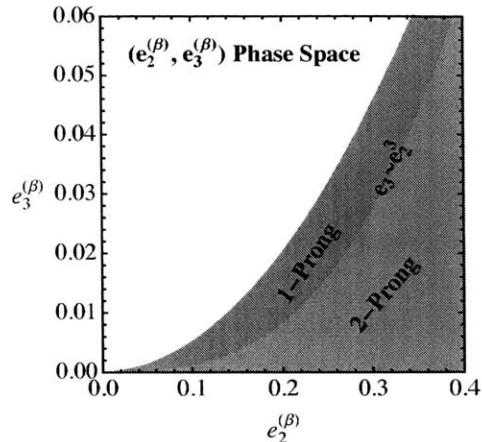
Figure 3-2: Phase space defined by the measurement of the energy correlation functions e_2 and e_3. The phase space is divided into 1- and 2-prong regions with a boundary corresponding to the curve $e_3 \sim (e_2)^3$.

subjets with characteristic p_T fraction $z_s \ll 1$. For color-singlet jets, like boosted Z bosons, this global soft radiation contribution comes purely from initial state radiation (ISR).[10] Finally, there is radiation from the dipole formed from the two subjets (called "collinear-soft" radiation), with characteristic angle R_{12} from the subjets, and with p_T fraction z_{cs}. The effective theory of this phase space region for the observable N-jettiness [104] was studied in Ref. [105].

We now consider the power counting of $e_2^{(\beta)}$ and $e_3^{(\beta)}$ for 2-prong jets. By the definition of this region of phase space, the hard splitting sets the value of $e_2^{(\beta)}$. That is, we have $e_2^{(\beta)} \sim R_{12}^{\beta}$, with all other contributions suppressed. For $e_3^{(\beta)}$, it is clear that the leading contributions must arise from correlations between the two hard subjets with either the global soft, collinear or collinear-soft modes. The scaling of these different contributions to $e_3^{(\beta)}$ is given in Table 3.2, from which we find that the scaling of the two- and three-point

---

[10]While this background is to a certain extent irreducible, given the important feature of ISR is its color-uncorrelated and soft nature, many of our observations about the effects of pile-up in Sec. 3.3.2 will be applicable to ISR.

modes	$e_3^{(\beta)}$
$C_1 C_2\, S$	$R_{12}^{\beta} z_s$
$C_1 C_2\, C$	$R_{12}^{2\beta} R_{cc}^{\beta}$
$C_1 C_2\, C_s$	$R_{12}^{3\beta} z_{cs}$

Table 3.2: Scaling of the contributions from global soft (S), collinear (C), and collinear-soft (C_s) radiation correlated with the two hard subjets (denoted by C_1 and C_2) in 2-prong jets to $e_3^{(\beta)}$ from the different possible configurations.

energy correlation functions for 2-pronged jets is

$$e_2^{(\beta)} \sim R_{12}^{\beta}, \tag{3.14}$$

$$e_3^{(\beta)} \sim R_{12}^{\beta} z_s + R_{12}^{2\beta} R_{cc}^{\beta} + R_{12}^{3\beta} z_{cs}. \tag{3.15}$$

There is no measurement performed to distinguish the three contributions to $e_3^{(\beta)}$ and so we must assume that they all scale equally.

This result is sufficient to set the relative scaling of $e_2^{(\beta)}$ and $e_3^{(\beta)}$. As we assume that the jet only has two hard subjets, we have that $z_{cs} \ll 1$ and so

$$(e_2)^3 \sim R_{12}^{3\beta} \gg R_{12}^{3\beta} z_{cs} \sim e_3, \tag{3.16}$$

which defines the 2-prong jet region of phase space as that for which $e_3 \ll (e_2)^3$. With this identification, note the scaling of the various modes:

$$R_{12}^{\beta} \sim e_2, \quad z_s \sim \frac{e_3}{e_2}, \quad R_{cc}^{\beta} \sim \frac{e_3}{(e_2)^2}, \quad z_{cs} \sim \frac{e_3}{(e_2)^3}. \tag{3.17}$$

While not important for our goals here, the fact that the energy correlation functions parametrically separate the scaling of the modes that contribute to the observables is vital for an effective theory analysis and calculability. This will be discussed in detail in Chap. 4, where an explicit calculation of the D_2 observable will be presented. Note that because e_2 is first non-zero at a lower order in perturbation theory than e_3, e_3 can be zero while e_2 is non-zero.

Therefore, this 2-prong region of phase space extends down to the kinematic limit of $e_3 = 0$, as shown in red in Fig. 3-2.

This power counting analysis, although very simple in nature, provides a powerful picture of the phase space defined by the measurement of e_2 and e_3, which is shown in Fig. 3-2. The 1- and 2-prong jets are defined to populate the phase space regions where

$$1\text{-prong jet: } (e_2)^3 \lesssim e_3 \lesssim (e_2)^2 \,,$$
$$2\text{-prong jet: } 0 < e_3 \ll (e_2)^3 \,.$$

Background QCD jets dominantly populate the 1-prong region of phase space, while signal boosted Z decays dominantly populate the 2-prong region of phase space. This has important consequences for the optimal discrimination observable.

An interesting observation about the boundary between 1- and 2-prong jets, defined by $e_3 \sim (e_2)^3$, is that it is approximately invariant to boosts along the jet direction. For a narrow jet, a boost along the jet direction by an amount γ scales p_Ts and angles as

$$p_T \to \gamma p_T \,, \qquad R \to \gamma^{-1} R \,. \tag{3.18}$$

Therefore, under a boost, $e_2^{(\beta)}$ and $e_3^{(\beta)}$ scale as

$$e_2^{(\beta)} \to \gamma^{-\beta} e_2^{(\beta)} \,, \qquad e_3^{(\beta)} \to \gamma^{-3\beta} e_3^{(\beta)} \,. \tag{3.19}$$

Thus, the boundary between 1- and 2-prong jets, where $e_3 \sim (e_2)^3$, is invariant to boosts along the jet direction. That is, under boosts, a jet will move along a contour of constant $e_3/(e_2)^3$ in the (e_2, e_3) plane.

The analysis presented in this section is also the initial step in establishing rigorous factorization theorems in the different regions of phase space, allowing for analytic resummation of the double differential cross section of e_2 and e_3. Such a factorization and resummation will be presented in Chap. 4, building on the analysis of this chapter.

Optimal Discrimination Observables

The fact that the signal and background regions of phase space are parametrically separated implies that from power counting alone, we can determine the optimal observable for separating signal from background. Because the boundary between the background-rich and signal-rich regions is $e_3 \sim (e_2)^3$, this suggests that the optimal observable for discriminating boosted Z bosons from QCD jets is[11]

$$D_2^{(\beta)} \equiv \frac{e_3^{(\beta)}}{\left(e_2^{(\beta)}\right)^3} . \qquad (3.20)$$

Signal jets will be characterized by a small value of $D_2^{(\beta)}$, while background jets will predominantly have large $D_2^{(\beta)}$. With this observable, parametrically there is no mixing of the signal-rich and background-rich regions. Contours of constant $D_2^{(\beta)}$ lie entirely in the signal or background region, as is shown schematically in Fig. 3-3. Determining the precise discrimination power of $D_2^{(\beta)}$ requires an understanding of the $\mathcal{O}(1)$ details of the distributions of signal and background, beyond any purely power counting analysis.

The observation that the scaling relation $e_3 \sim (e_2)^3$ is boost invariant provides further motivation for the variable $D_2^{(\beta)}$. Under boosts along the jet axis, jets can move along curves of constant $D_2^{(\beta)}$, but cannot cross the boundary between the 2-prong and 1-prong regions of phase space. This can be used to give a boost invariant definition of a 2-prong jet, as a jet with a small value of $D_2^{(\beta)}$, and a 1-prong jet, as a jet with large $D_2^{(\beta)}$.

Ref. [30] used the two- and three-point energy correlation functions in the combination

$$C_2^{(\beta)} \equiv \frac{e_3^{(\beta)}}{\left(e_2^{(\beta)}\right)^2} \qquad (3.21)$$

for boosted Z boson discrimination. From the power counting analysis in this section, this

---

[11]We thank Jesse Thaler for suggesting the notation "D" for these observables. Unlike $C_2^{(\beta)}$, whose name was motivated by its relation to the classic e^+e^- event shape parameter C, $D_2^{(\beta)}$ is not related to the D parameter.
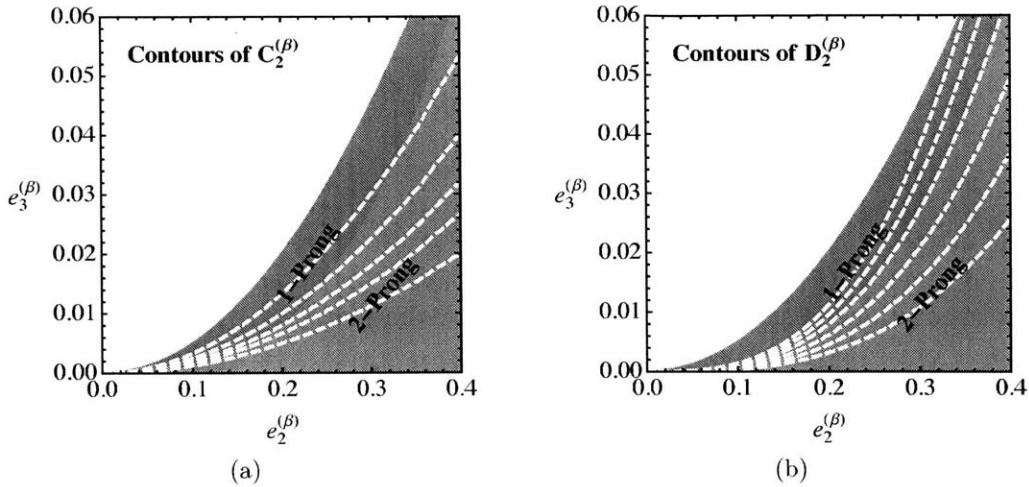
Figure 3-3: Contours of constant $C_2^{(\beta)}$ (left) and $D_2^{(\beta)}$ (right) in the phase space defined by $e_2^{(\beta)}, e_3^{(\beta)}$. The 1- and 2-prong regions of phase space are labeled, with their boundary corresponding to the curve $e_3 \sim (e_2)^3$.

variable is not a natural choice. In particular, contours of constant $C_2^{(\beta)}$ pass through both the 1-prong and 2-prong regions of phase space, mixing the signal and background for any value of $C_2^{(\beta)}$, as shown in Fig. 3-3. Therefore, from the power counting perspective, we would expect that $C_2^{(\beta)}$ is a poor boosted Z boson discriminating observable. Nevertheless, Ref. [30] found that with a tight jet mass cut, and in the absence of pile-up, $C_2^{(\beta)}$ is a powerful boosted Z discriminant. A mass cut constrains the phase space significantly, which we will discuss in detail in Sec. 3.3.1, allowing us to understand the result of Ref. [30]. Pile-up will be addressed in Sec. 3.3.2.

It is important to recall that while $e_2^{(\beta)}$ and $e_3^{(\beta)}$ are IRC safe observables, so that their phase space can be analyzed with power counting techniques, ratios of IRC safe observables are not in general IRC safe [106, 88, 89]. The observables $C_2^{(\beta)}$ and $D_2^{(\beta)}$ are however Sudakov safe [88, 89], and therefore can be reliably studied with Monte Carlo simulation without applying any form of additional cut, such as a jet mass cut, on the phase space.

modes	$\tau_1^{(\beta)}$	$\tau_2^{(\beta)}$
CCC	R_{cc}^{β}	R_{cc}^{β}
CCS	$R_{cc}^{\beta} + z_s$	$R_{cc}^{\beta} + z_s$
CSS	z_s	z_s
SSS	z_s	z_s

Table 3.3: Scaling of the contributions of 1-prong jets to $\tau_1^{(\beta)}$ and $\tau_2^{(\beta)}$ from the different possible configurations of soft (S) and collinear (C) radiation.

Contrasting with N-subjettiness

At this point, it is interesting to apply the power counting analysis to other observables for boosted Z discrimination and see what conclusions can be made. For concreteness, we will contrast the energy correlation functions with the N-subjettiness observables $\tau_1^{(\beta)}$ and $\tau_2^{(\beta)}$, defined as

$$\tau_N^{(\beta)} = \frac{1}{p_{TJ}} \sum_{1 \leq i \leq n_J} p_{Ti} \min \left\{ R_{i1}^{\beta}, \ldots, R_{iN}^{\beta} \right\} . \tag{3.22}$$

As with the energy correlation functions, we will consider $\tau_1^{(\beta)}$ and $\tau_2^{(\beta)}$ as measured on 1-prong and 2-prong jets and determine the regions of phase space where background and signal jets populate. This can then be used to determine the optimal observable for boosted Z discrimination from the N-subjettiness observables. We use the same notation for the scalings of the modes as in Sec. 3.3.1.

Starting with 1-prong jets, and repeating the analysis of Sec. 3.3.1, we find the dominant contributions to $\tau_1^{(\beta)}$ and $\tau_2^{(\beta)}$ as given in Table 3.3. For the configuration of two collinear particles and a soft particle (CCS), $\tau_2^{(\beta)}$ is either dominated by z_s or by R_{cc}^{β}. In this configuration, the two subjet axes can either lie on the two collinear particles or one axis can be on a collinear particle and the other on a soft particle. Importantly, the measurement of $\tau_1^{(\beta)}$ and $\tau_2^{(\beta)}$ cannot distinguish these two possibilities and therefore cannot determine if the second axis in the 1-prong jet is at a small or large angle with respect to the first.[12] With

---

[12]For this reason, soft and collinear contributions to $\tau_2^{(\beta)}$ on 1-prong jets do not factorize and therefore cannot be computed in SCET.

modes	$\tau_2^{(\beta)}$
$C_1 C_2 \, S$	z_s
$C_1 C_2 \, C$	R_{cc}^{β}
$C_1 C_2 \, C_s$	$R_{12}^{\beta} z_{cs}$

Table 3.4: Scaling of the contributions from global soft (S), collinear (C), and collinear-soft (C_s) radiation correlated with the two hard subjets (denoted by C_1 and C_2) in 2-prong jets to $\tau_2^{(\beta)}$ from the different possible configurations.

either configuration, $\tau_1^{(\beta)}$ and $\tau_2^{(\beta)}$ scale as

$$\tau_1^{(\beta)} \sim R_{cc}^{\beta} + z_s \,, \tag{3.23}$$

$$\tau_2^{(\beta)} \sim R_{cc}^{\beta} + z_s \,. \tag{3.24}$$

That is, for 1-prong jets, $\tau_1^{(\beta)} \sim \tau_2^{(\beta)}$.

For 2-prong jets, $\tau_1^{(\beta)}$ is dominated by the hard splitting, as was the case with the two-point energy correlation function, hence $\tau_1^{(\beta)} \sim R_{12}^{\beta}$. For $\tau_2^{(\beta)}$, the two axes lie along the two hard prongs, so, just like with the three-point energy correlation function, $\tau_2^{(\beta)}$ is set by the radiation about those two hard prongs: global soft, collinear, or collinear soft. Table 3.4 lists the contributions to $\tau_2^{(\beta)}$ from each of these modes, leading to the scaling

$$\tau_1^{(\beta)} \sim R_{12}^{\beta} \,, \tag{3.25}$$

$$\tau_2^{(\beta)} \sim z_s + R_{cc}^{\beta} + R_{12}^{\beta} z_{cs} \,. \tag{3.26}$$

Demanding that the jet only has two hard prongs implies that $\tau_2^{(\beta)} \sim z_s \sim R_{cc}^{\beta} \sim R_{12}^{\beta} z_{cs} \ll R_{12}^{\beta} \sim \tau_1^{(\beta)}$, but no other conclusions can be made from power counting alone. Unlike the well-defined division of phase space by the energy correlation functions, N-subjettiness has a much weaker division of

$$\text{1-prong jet: } \tau_2^{(\beta)} \sim \tau_1^{(\beta)} \,,$$

$$\text{2-prong jet: } \tau_2^{(\beta)} \ll \tau_1^{(\beta)} \,.$$
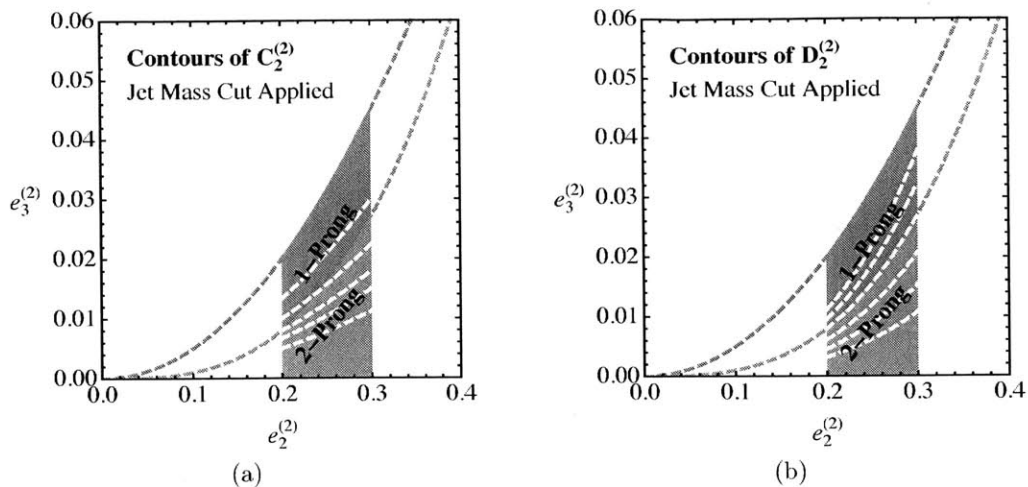
Figure 3-4: Phase space defined by the energy correlation functions $e_2^{(2)}, e_3^{(2)}$ in the presence of a mass cut. Contours of constant $C_2^{(2)}$ (left) and $D_2^{(2)}$ (right) are shown for reference.

This does suggest, however, that the optimal discrimination variable using N-subjettiness is $\tau_{2,1}^{(\beta)} \equiv \tau_2^{(\beta)} / \tau_1^{(\beta)}$, which is what is widely used experimentally. Nevertheless, the weaker phase space separation of N-subjettiness compared with that for the energy correlation functions would naïvely imply that e_2 and e_3 provides better discrimination than $\tau_1^{(\beta)}$ and $\tau_2^{(\beta)}$; however, this statement requires an understanding of $\mathcal{O}(1)$ numbers, which is beyond the scope of a power counting analysis.

Effect of a Mass Cut

In an experimental application of $D_2^{(\beta)}$ to boosted Z discrimination, a mass cut is performed on the jet around the mass of the Z boson. In addition to removing a large fraction of the background, this cut also guarantees that the identified jets are actually generated from boosted Z decays. To fully understand the effect of the mass cut on the phase space requires analyzing the three-dimensional phase space of the mass, e_2, and e_3. While complete, this full analysis would be distracting to the physics points that we wish to make in this section, and the impact of the mass cut can be understood without performing this analysis. For $\beta = 2$, the two-point energy correlation function is simply related to the jet mass m at fixed

jet p_T:

$$e_2^{(2)} \simeq \frac{m^2}{p_T^2} \ , \tag{3.27}$$

for central jets assuming that $m \ll p_T$ and up to overall factors of order 1. Therefore, a cut on the jet mass is a cut on $e_2^{(2)}$. In this section, we will begin by discussing the simpler case of $\beta = 2$, and then proceed to comment on the effect of a mass cut for general β.

The phase space in the $e_2^{(2)}, e_3^{(2)}$ plane with the jet mass constrained to a window, and for some finite range of jet p_T is shown schematically in Fig. 3-4. Jets of a given mass can have that mass generated either by substantial soft radiation (for 1-prong jets) or by a hard splitting in the jet (a 2-prong jet), and so we want a discrimination observable that separates these two regions cleanly. The boundary between the 1-prong and 2-prong jet regions is still defined by $e_3^{(2)} \sim (e_2^{(2)})^3$, and so we expect $D_2^{(2)}$ to be the most powerful discriminant. However, by making a mass cut, the region of phase space at small masses, dominated by 1-prong jets, is removed. Therefore, the fact that contours of the observable $C_2^{(2)}$ mix both 1- and 2-prong jets is much less of an issue. Except at very high signal efficiencies, when one is sensitive to the functional form of the boundary between the signal and background regions, the discrimination performance of $C_2^{(2)}$ should be similar to that of $D_2^{(2)}$ when a tight mass cut is imposed. Indeed, in a sufficiently narrow window, any variable of the form $e_3^{(2)}/(e_2^{(2)})^n$, would provide reasonable discrimination, with all the discrimination power coming from $e_3^{(2)}$ alone. However, $D_2^{(2)}$ has the advantage that its discrimination power does not suffer from significant dependence on the value of the lower mass cut.

While a lower mass cut is important for removing 1-prong background jets, an upper mass cut is also necessary for powerful discrimination. The mass distribution of QCD jets has a long tail extending to masses of order the p_T of the jet. For these jets, the mass is generated by an honest hard splitting, and so these background jets look exactly like the signal from their substructure. While the cross section for these high mass QCD jets is suppressed by α_s, they can still be a significant background and therefore should be removed.

Let's now consider the general β case. We will first consider the effect of a mass cut in
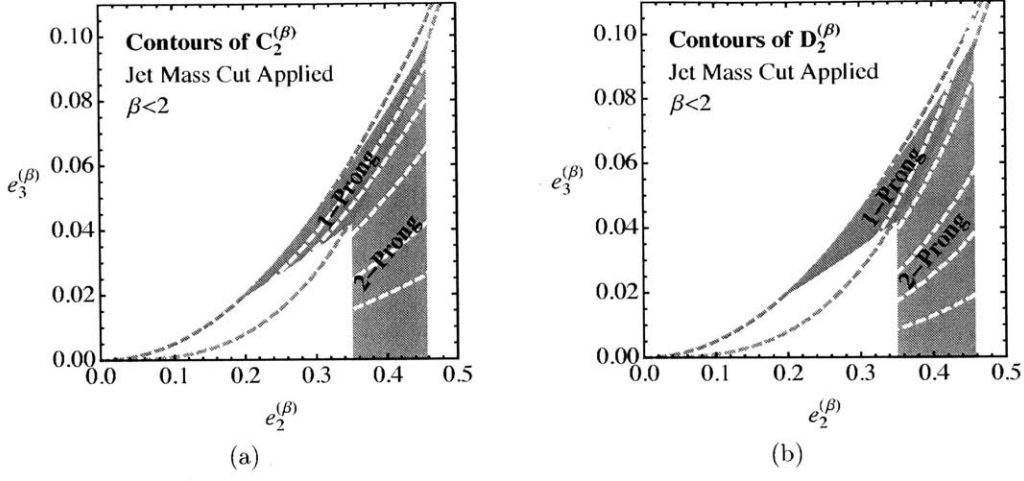
Figure 3-5: Phase space defined by the energy correlation functions $e_2^{(\beta)}, e_3^{(\beta)}$, for $\beta < 2$, in the presence of a mass cut. Contours of constant $C_2^{(2)}$ (left) and $D_2^{(2)}$ (right) are shown for reference.

the 2-prong region of the $e_2^{(\beta)}, e_3^{(\beta)}$ plane. Recall that in this region of phase space

$$e_2^{(\beta)} \sim R_{12}^\beta , \tag{3.28}$$

where R_{12} is the angle between the hard subjets. Therefore, in this region of phase space $e_2^{(\beta)}$ is simply related to the mass:

$$e_2^{(\beta)} \sim \left(\frac{m^2}{p_T^2} \right)^{\beta/2} . \tag{3.29}$$

A cut on the jet mass is therefore equivalent to an appropriate cut on $e_2^{(\beta)}$ for 2-prong jets.

A mass cut in the 1-prong region of phase space is more subtle, as the dominant contributing mode to $e_2^{(\beta)}$ changes throughout the phase space. Recall that in this region, $e_2^{(\beta)}$ has the scalings

$$e_2^{(\beta)} \sim R_{cc}^\beta + z_s . \tag{3.30}$$

82

while

$$e_2^{(2)} \sim \frac{m^2}{p_T^2} \sim R_{cc}^2 + z_s \,. \tag{3.31}$$

While the soft contributions have the same scaling for both variables, the collinear contributions do not. There are two possibilities as for the relative scalings of $e_2^{(\beta)}$ and the mass: if soft emissions do not contribute, then

$$e_2^{(\beta)} \sim \left(\frac{m^2}{p_T^2}\right)^{\beta/2} \,, \tag{3.32}$$

which matches onto the relative scaling in the 2-prong region of phase space. If instead soft emissions do contribute, then

$$e_2^{(\beta)} \sim \frac{m^2}{p_T^2} \,, \tag{3.33}$$

which defines the upper boundary of the 1-prong phase space. These phase space boundaries for jets on which two two-point energy correlation functions with different angular exponents (or recoil-free angularities [33]) are measured is discussed in detail in Ref. [10].

Depending on whether β is less than or greater than 2, the mass cut manifests itself differently. For $\beta < 2$, note that from Eqs. (3.32) and (3.33), $e_2^{(\beta)} > e_2^{(2)}$ in the two-prong region, and so smaller values of $e_2^{(\beta)}$ can correspond to the same mass. Conversely, for $\beta > 2$, $e_2^{(\beta)} < e_2^{(2)}$ and so larger values of $e_2^{(\beta)}$ can correspond to the same mass. The effect of a mass cut on the allowed phase space for $\beta < 2$ is illustrated schematically in Fig. 3-4. Because in this case small values of $e_2^{(\beta)}$ can satisfy the mass cut, contours of $C_2^{(\beta)}$ can pass through the background region of phase space and significantly reduce the discrimination power. Again, because it respects the parametric scaling of the phase space boundaries, we expect the discrimination power of $D_2^{(\beta)}$ to be more robust as β decreases from 2. However, the precise discrimination power depends on understanding the $\mathcal{O}(1)$ region around the 1-prong and 2-prong jet boundary as β moves away from 2. This observation also explains why Ref. [30] found that the optimal choice for boosted Z boson discrimination using $C_2^{(\beta)}$ with a tight mass cut was $\beta \simeq 2$.

Summary of Power Counting Predictions

Here, we summarize the main predictions from our power counting analysis of boosted Z discrimination, before a Monte Carlo study in Sec. 3.3.1. We have:

- The parametric scaling of the boundary between 1-prong and 2-prong jets in the (e_2, e_3) phase space is $e_3 \sim (e_2)^3$. Therefore, $D_2^{(\beta)}$ should be a more powerful discrimination observable than $C_2^{(\beta)}$ because contours of constant $D_2^{(\beta)}$ do not mix signal and background regions, while contours of $C_2^{(\beta)}$ do mix signal and background regions.

- When a mass cut is imposed on the jet, $C_2^{(\beta)}$ should have similar discrimination power to $D_2^{(\beta)}$, for $\beta \simeq 2$, except at high signal efficiency when the observable is sensitive to the boundary between the signal and background regions. At high signal efficiency, $D_2^{(\beta)}$ should be a slightly better discriminant than $C_2^{(\beta)}$ for $\beta \simeq 2$.

- The discrimination power of $C_2^{(\beta)}$ should decrease substantially as β decreases from 2 when there is a mass cut on the jets. By contrast, the discrimination power of $D_2^{(\beta)}$ should be more robust as β decreases from 2.

- The power counting predictions stated above should be robust to Monte Carlo tuning and reproduced by any Monte Carlo simulation, e.g. HERWIG++ or PYTHIA 8, since they are determined by parametric scaling of QCD dynamics.

Monte Carlo Analysis

To test these predictions, we will study the different ratio observables formed from $e_2^{(\beta)}$ and $e_3^{(\beta)}$ in Monte Carlo simulation. We generated background QCD jets from $pp \to Zj$ events, with the Z decaying leptonically, and boosted Z decays from $pp \to ZZ$ events, with one Z decaying leptonically, and the other to quarks. Events were generated with MADGRAPH5 2.1.2 [107] at the 8 TeV LHC, and showered with either PYTHIA 8.183 [108, 109] or HERWIG++ 2.6.3 [110, 111, 112, 113], to test the robustness of our predictions to the details of the Monte Carlo generator. Anti-k_T [94] jets with radius $R = 1.0$ and $p_T > 400$
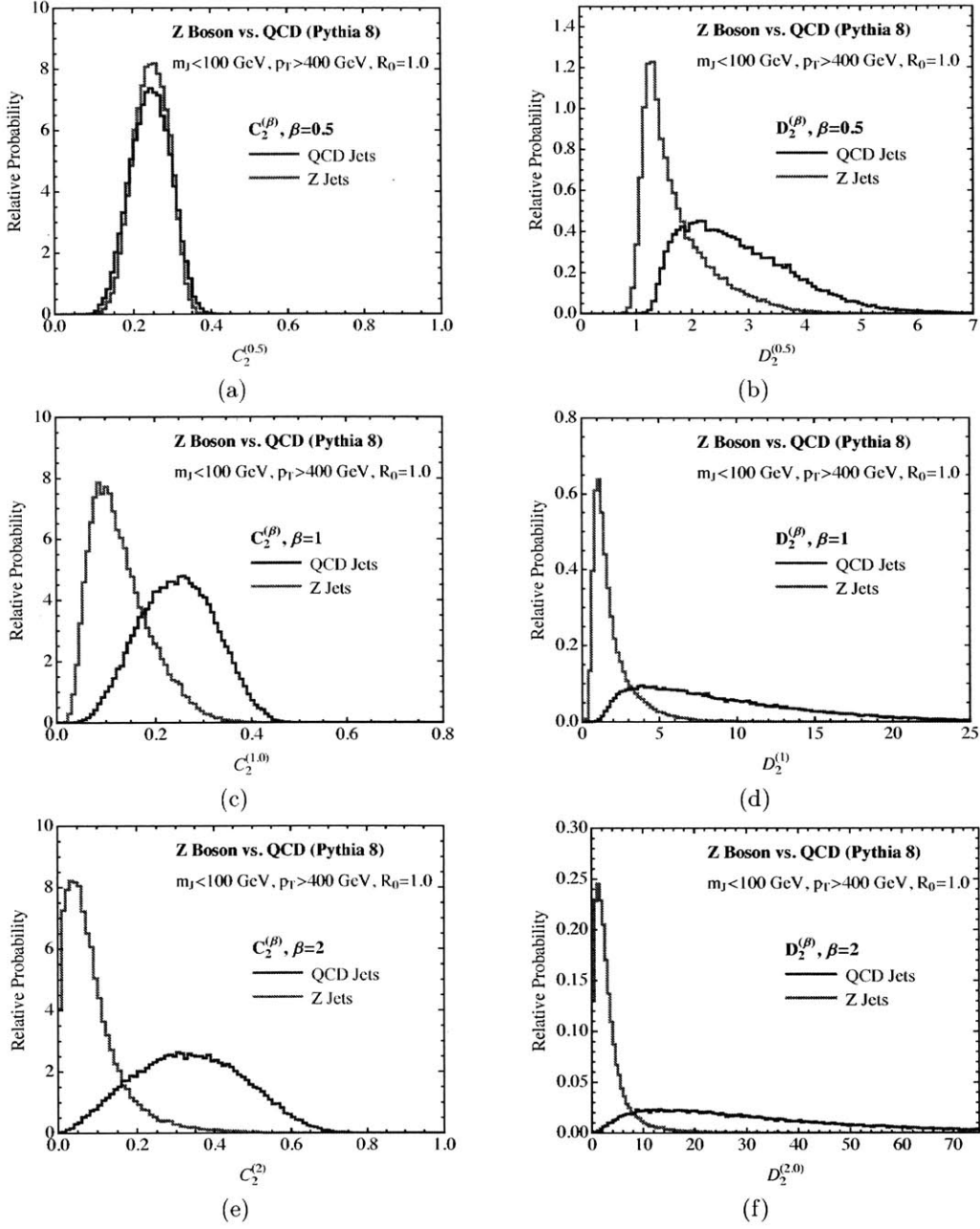
Figure 3-6: Signal and background distributions for the ratio observables $C_2^{(\beta)}$ (left) and $D_2^{(\beta)}$ (right) for $\beta = 0.5, 1, 2$ from the MADGRAPH5 and PYTHIA 8 samples. No lower mass cut on the jets is applied but we take $m_J < 100$ GeV.
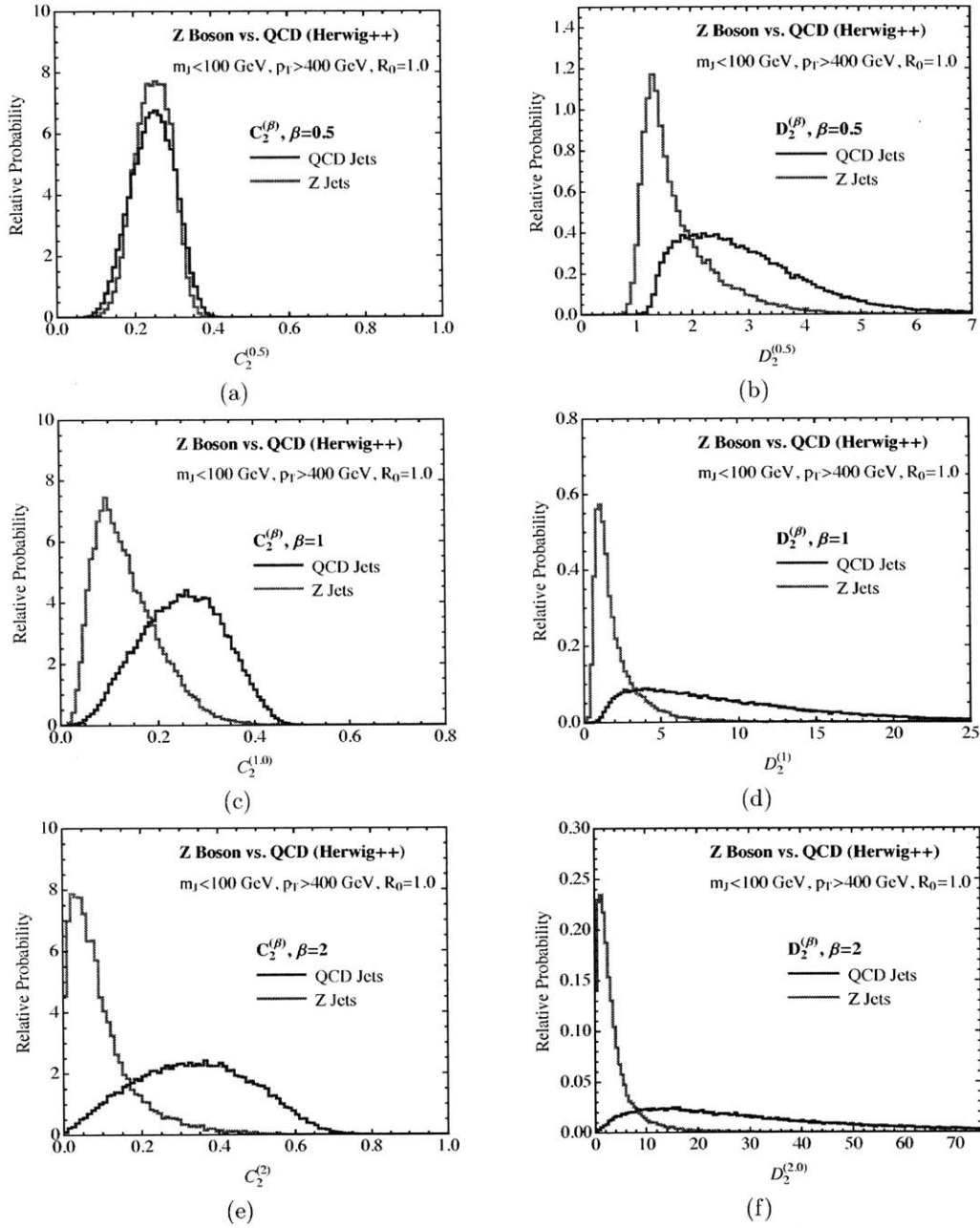
Figure 3-7: Same plots as in Fig. 3-6, from the HERWIG++ samples.
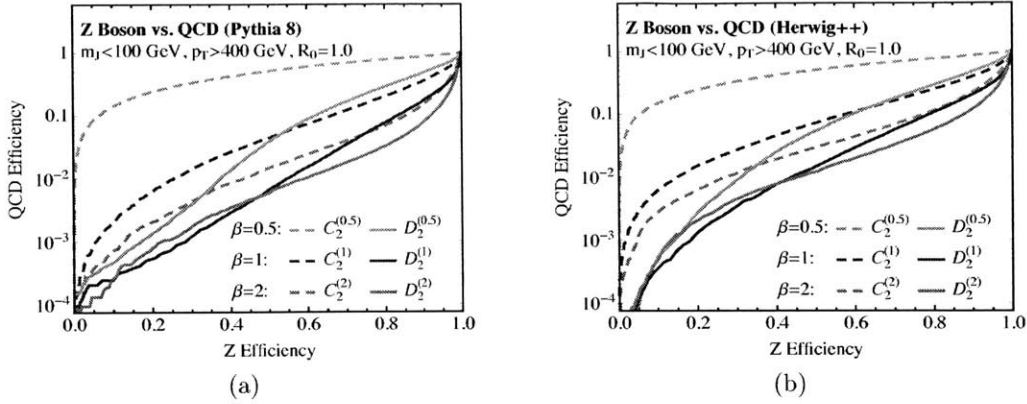
Figure 3-8: Signal vs. background efficiency curves (ROC curves) for $C_2^{(\beta)}$ and $D_2^{(\beta)}$ for $\beta = 0.5, 1, 2$ for jets with $m_J < 100$ GeV, showered with PYTHIA 8 (left) and HERWIG++ (right). Power counting predictions for the behavior of the ROC curves are robustly reproduced by both Monte Carlo generators.

GeV were clustered in FASTJET 3.0.3 [100] using the Winner Take All (WTA) recombination scheme [33, 89]. The energy correlation functions and N-subjettiness ratio observables were calculated using the EnergyCorrelator and Nsubjettiness FASTJET CONTRIBs [100, 101].

We first compare the discrimination power of $C_2^{(\beta)}$ to $D_2^{(\beta)}$ with no lower mass cut on the jets for several values of the angular exponent. We require that $m_J < 100$ GeV which removes a significant fraction of QCD jets that have honest 2-prong structure. Therefore, we are testing the power of $C_2^{(\beta)}$ and $D_2^{(\beta)}$ to discriminate between 1-prong and 2-prong jets. In Fig. 3-6, we show the raw distributions of $C_2^{(\beta)}$ and $D_2^{(\beta)}$ measured on signal and background for $\beta = 0.5, 1, 2$. Especially at small β, $D_2^{(\beta)}$ is much more efficient at separating boosted Zs from QCD jets than is $C_2^{(\beta)}$. This is exactly as predicted by the power counting, because $C_2^{(\beta)}$ mixes the signal and background regions of phase space, an effect that is magnified at smaller β. The discrimination power is quantified in Fig. 3-8 where we show the signal vs. background efficiency curves (ROC curves) for the three choices of β for $C_2^{(\beta)}$ and $D_2^{(\beta)}$. At low signal efficiency, every $D_2^{(\beta)}$ is a better discriminant than any $C_2^{(\beta)}$, and the performance of $D_2^{(\beta)}$ is much more stable as a function of β than $C_2^{(\beta)}$.

In the presence of a narrow mass cut window, the power counting analysis of Sec. 3.3.1

87

predicted that for β near 2, the discrimination power of $C_2^{(\beta)}$ and $D_2^{(\beta)}$ should be comparable except at high signal efficiency when $D_2^{(\beta)}$ should be more discriminating. To show that this is borne out in Monte Carlo, in Fig. 3-9 we first plot the rejection efficiency of $C_2^{(1.7)}$ and $D_2^{(1.7)}$ at 90% signal efficiency, as a function of the lower mass cut on the jets.[13] When the lower mass cut is near zero, $D_2^{(1.7)}$ is significantly more efficient at rejecting QCD background than is $C_2^{(1.7)}$, as observed earlier. As the lower mass cut increases, however, the difference in discrimination power between the two observables decreases in both PYTHIA 8 and HER-WIG++ Monte Carlos. This dependence on the lower mass cut shows that $D_2^{(\beta)}$ captures the correct underlying physics of the $(e_2^{(\beta)}, e_3^{(\beta)})$ phase space, while $C_2^{(\beta)}$ does not. The light QCD jets that are added as the mass cut is lowered should be rejected by a variable that partitions the phase space into regions of 1-prong and 2-prong jets, increasing the observed rejection efficiency. This is true for $D_2^{(\beta)}$; however, exactly the opposite is true for $C_2^{(\beta)}$.

We now study in more detail the case in which we have constrained the jet mass to lie in the tight mass cut window of $80 < m_J < 100$ GeV. Over the whole signal efficiency range, $C_2^{(1.7)}$ and $D_2^{(1.7)}$ have nearly identical ROC curves in both PYTHIA 8 and HERWIG++, as exhibited in Fig. 3-10. However, focusing in on the high signal efficiency region, we see that indeed $D_2^{(1.7)}$ has a slightly better rejection rate than $C_2^{(1.7)}$. This behavior is manifest in both PYTHIA 8 and HERWIG++, showing that this prediction from the power counting analysis of Sec. 3.3.1 is robust to the precise details of the parton shower in the Monte Carlo generator. This should be contrasted with the actual numerical value of the QCD rejection, which depends on the generator. For $\beta \simeq 2$ with a tight mass cut window of $80 < m_J < 100$ GeV, any discriminating variable of the form $e_3^{(\beta)}/e_2^{(\beta)n}$, for $n > 0$, provides reasonable discrimination power. The jet mass cut fixes $e_2^{(2)}$ to a narrow window, and all discrimination power comes from $e_3^{(2)}$ alone. This demonstrates why Ref. [30] observed near-optimal discrimination power using $C_2^{(2)}$, with $80 < m_J < 100$ GeV.

The final power counting prediction for the behavior of the observables was that the discrimination power of $D_2^{(\beta)}$ should be much more robust than $C_2^{(\beta)}$ as β decreases from 2, with a mass cut on the jets. This behavior is reproduced in Monte Carlo, as shown in Fig. 3-

---

[13]We use $\beta = 1.7$ as this value was shown in Ref. [30] to be the optimal choice for boosted Z identification.
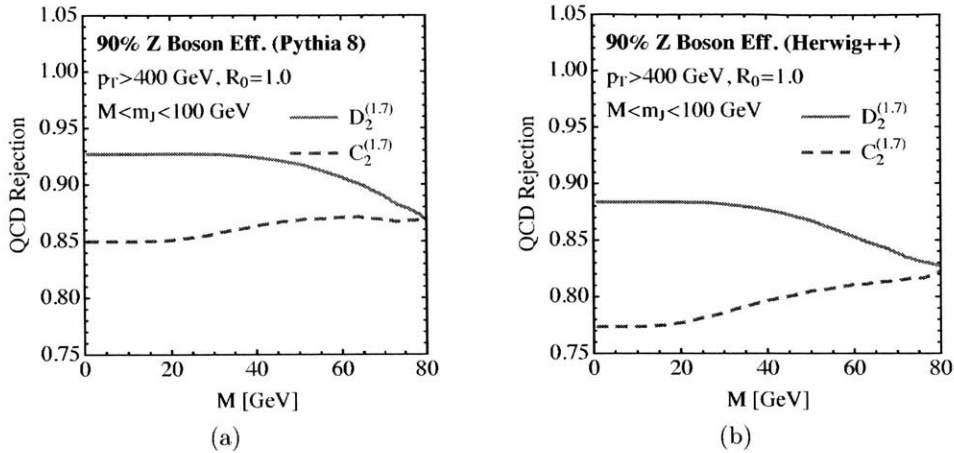
Figure 3-9: QCD rejection efficiency at 90% signal efficiency as a function of the lower mass cut, as predicted by PYTHIA 8 (left), and HERWIG++ (right). The plots compare the efficiencies of $C_2^{(1.7)}$ and $D_2^{(1.7)}$.

10, where we have plotted the QCD rejection efficiency at 50% signal efficiency as a function of β. We have also included in these plots the N-subjettiness ratio $\tau_{2,1}^{(\beta)}$ for comparison. As $\beta \to 0$, the discrimination power of both $C_2^{(\beta)}$ and $D_2^{(\beta)}$ decrease; however, $D_2^{(\beta)}$ maintains high discrimination power to much smaller values of β than $C_2^{(\beta)}$. Nevertheless, note that as $\beta \to 0$, $\tau_{2,1}^{(\beta)}$, while not the optimal discrimination observable, has even more robust discrimination power than $D_2^{(\beta)}$. An understanding of this behavior requires an analyses of $\mathcal{O}(1)$ numbers, which is beyond what power counting alone can predict.

Thus, we see that all the power counting predictions are realized in both Monte Carlo simulations, demonstrating that parametric scalings are indeed determining the behavior of the substructure observables. We emphasize that the level of agreement between the Monte Carlo generators for the power counting predictions is quite remarkable, given that numerical values for rejection or acceptance efficiencies, for example, do not agree particularly well between the generators. Power counting has allowed us to identify the robust predictions of perturbative QCD.

89
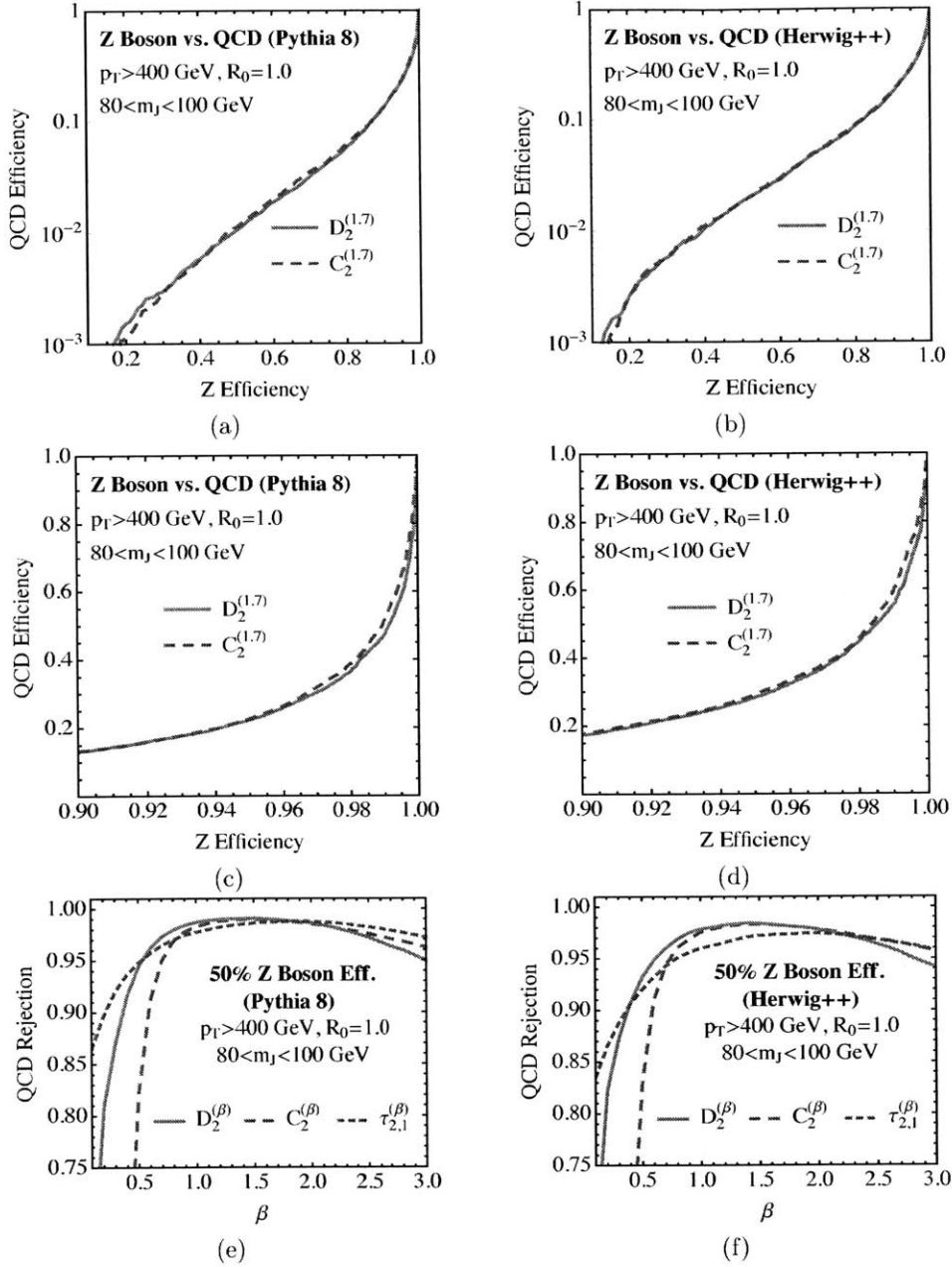
Figure 3-10: Comparison of $C_2^{(\beta)}$ and $D_2^{(\beta)}$ in the presence of a tight mass cut, $80 < m_J < 100$ GeV, with the PYTHIA 8 (left) and HERWIG++ (right) samples. ROC curves for $C_2^{(1.7)}$ and $D_2^{(1.7)}$ demonstrate that with a tight mass cut, both observables perform comparably over a large range of signal efficiencies (top), with $D_2^{(1.7)}$ performing slightly better at high signal efficiencies (middle), behavior which is reproduced by both Monte Carlo generators. The QCD rejection rate at 50% signal efficiency as a function of β is shown at bottom for $C_2^{(\beta)}$, $D_2^{(\beta)}$ and $\tau_{2,1}^{(\beta)}$.

90

3.3.2 Including Pile-Up

The power counting analysis of the previous section included only perturbative radiation. At a high luminosity hadron collider such as the LHC, also important is the effect of multiple proton collisions per bunch crossing, referred to as pile-up. Pile-up radiation is uncorrelated with the hard scattering event, and as such, has an energy scale that is independent of the hard parton collision energy. Thus, pile-up can produce a significant amount of contaminating radiation in the event and substantially change jet p_Ts, masses, or observables from their perturbative values. An important problem in jet substructure is both to define observables that are less sensitive to the effects of pile-up, as well as to remove or "groom", to the greatest extent possible, radiation in a jet or event that most likely is from pile-up. Several methods for jet grooming and pile-up subtraction have been presented [114, 115, 116, 117, 106, 118, 40, 119, 120, 121], and are used by the experiments [63, 66, 61, 62, 51], but we will not consider them here.[14] Instead, we will demonstrate that power counting can be used to understand the effect of pile-up radiation on the $(e_2^{(\beta)}, e_3^{(\beta)})$ phase space, and therefore on signal and background distributions for observables formed from the energy correlation functions. We envision that similar techniques could be used to develop jet substructure variables with improved resilience to pile-up, but in this chapter we will restrict ourselves to an understanding of the behavior of $C_2^{(\beta)}$ and $D_2^{(\beta)}$.[15]

To incorporate pile-up radiation into the power counting analysis, we must make some simplifying assumptions. Because pile-up is independent of the hard scattering event, we will assume that pile-up radiation is uniformly distributed over the jet area.[16] This assumption essentially defines pile-up as another soft mode in the jet, with all angles associated with

---

[14]The effects of jet grooming techniques can be understood using power counting techniques, and have been considered in [90].

[15]While the following analysis is quite general, it is restricted to recoil-free observables defined with a recoil-free jet algorithm, as used in this chapter. In the case of a recoil sensitive observable, there is a non-linear response to pile-up due to the displacement of soft and collinear modes with respect to the jet axis. In this case the power counting analysis described here does not apply directly, and a more thorough analysis is required.

[16]This model of pile-up would be removed by area subtraction [106]. However, this would also remove perturbative soft radiation depending on the region of phase space. This could be studied in detail using power counting.

pile-up scaling as $\mathcal{O}(1)$. We will denote the p_T fraction of pile-up radiation in the jet as

$$z_{pu} \equiv \frac{p_{Tpu}}{p_{TJ}} \, . \tag{3.34}$$

No assumption of the relative size of the perturbative soft radiation energy fraction z_s with respect to z_{pu} is made at this point, and indeed the impact of pile-up on the phase space will depend on this relation.

Assuming only that the pile-up p_T fraction $z_{pu} \ll 1$, the two- and three-point correlation functions for 1-prong jets have the scaling

$$e_2^{(\beta)} \sim R_{cc}^\beta + z_s + z_{pu} \, , \tag{3.35}$$

$$e_3^{(\beta)} \sim R_{cc}^{3\beta} + z_s^2 + R_{cc}^\beta z_s + z_{pu}^2 + R_{cc}^\beta z_{pu} \, . \tag{3.36}$$

For 2-prong jets, the correlation functions have the scaling

$$e_2^{(\beta)} \sim R_{12}^\beta + z_{pu} \, , \tag{3.37}$$

$$e_3^{(\beta)} \sim R_{12}^\beta z_s + R_{12}^{2\beta} R_{cc}^\beta + R_{12}^{3\beta} z_{cs} + R_{12}^\beta z_{pu} + z_{pu}^2 \, . \tag{3.38}$$

From these scalings, we will be able to understand how pile-up radiation impacts jets in different regions of phase space, and hence the distributions in $C_2^{(\beta)}$ and $D_2^{(\beta)}$. Note that z_{pu} is a fixed quantity measuring the fraction of pile-up radiation in the jet, and unlike the scalings for the soft, collinear and collinear-soft modes, its scaling is constant throughout the phase space. To understand the impact of the pile-up radiation on different regions of phase space, we will therefore need to understand how the values of $e_2^{(\beta)}$ and $e_3^{(\beta)}$ are modified by z_{pu}, depending on the different scalings of the contributing modes.

We begin the study of the phase space at small $e_2^{(\beta)}$. In the limit when $z_{pu} \gg z_s$, pile-up dominates the structure of the jet. In this limit, both 1-prong and 2-prong jets are forced into the region of phase space where $e_3^{(\beta)} \sim (e_2^{(\beta)})^2$. Note however that the scaling of the upper boundary of the phase space is robust. We must assume, as we will in what follows,
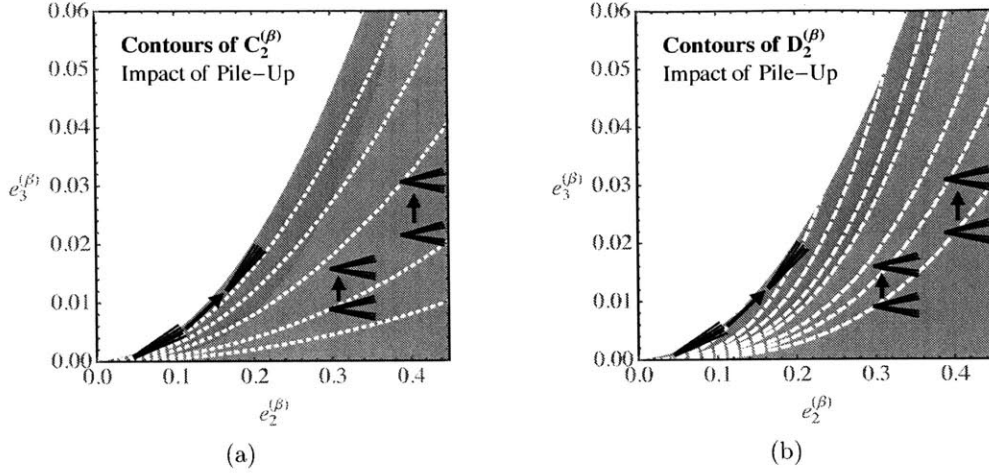
Figure 3-11: Illustration of the effect of pile-up on the e_2, e_3 phase space. The 1- and 2-prong regions of phase space are denoted by blue or red, respectively, and the arrows show the direction that the jets move in the phase space with the addition of pile-up. Contours of constant C_2 (left) and D_2 (right) are shown for reference.

that the value of z_{pu} is such that this region does not extend far into the phase space, or else the energy correlation functions cannot be used to discriminate 1- and 2-prong jets, as their structure is completely dominated by pile-up radiation.

Moving to slightly larger values of $e_2^{(\beta)}$, we encounter a region dominated by 1-prong background jets, where $z_{pu} \sim z_s$. Under the addition of pile-up radiation, the two- and three-point correlation functions for 1-prong jets are modified as

$$e_2^{(\beta)} \to e_2^{(\beta)} + z_{pu}\,, \tag{3.39}$$

$$e_3^{(\beta)} \to e_3^{(\beta)} + e_2^{(\beta)} z_{pu} + z_{pu}^2\,. \tag{3.40}$$

For $z_{pu} \sim z_s$, the addition of pile-up radiation therefore pushes all 1-prong jets towards the boundary $e_3^{(\beta)} \sim (e_2^{(\beta)})^2$. Jets that already satisfy this scaling, maintain it under the addition of pile-up, but move to larger values of $e_2^{(\beta)}$. This behavior is illustrated in Fig. 3-11, and will imply a very different behavior for the distributions of $C_2^{(\beta)}$ and $D_2^{(\beta)}$ for background.

At larger values of $e_2^{(\beta)}$, populated primarily by jets with two hard prongs, pile-up is a

93

power-suppressed contribution to e_2, but still contributes to e_3. That is, pile-up affects the two- and three-point correlation functions measured on 2-prong jets as

$$e_2^{(\beta)} \to e_2^{(\beta)} \,, \tag{3.41}$$

$$e_3^{(\beta)} \to e_3^{(\beta)} + e_2^{(\beta)} z_{pu} \,. \tag{3.42}$$

Therefore, pile-up shifts 2-prong jets vertically in the (e_2, e_3) phase space plane by an amount proportional to the perturbative value of $e_2^{(\beta)}$. This behavior is illustrated schematically in Fig. 3-11. At even larger values of $e_2^{(\beta)}$, we enter a regime where $z_{pu} \ll z_s$. Here the scale of the pile-up radiation is parametrically smaller than the soft perturbative radiation, and so to leading power pile-up can be ignored.

We will now use this understanding of the effect of pile-up radiation on different regions of the (e_2, e_3) phase space to understand its impact on the distributions for the observables $C_2^{(\beta)}$ and $D_2^{(\beta)}$ for both the signal and background. We begin by discussing the impact of pile-up radiation on the background distribution of $D_2^{(\beta)}$. Recall that at small $e_2^{(\beta)}$, both $e_2^{(\beta)}$ and $e_3^{(\beta)}$ are shifted by the addition of the pile-up radiation, but maintain the parametric scaling $e_3^{(\beta)} \sim (e_2^{(\beta)})^2$. This has an interesting effect on the $D_2^{(\beta)}$ distribution due to the fact that the functional form of the contours, which are cubic, does not match the quadratic scaling of the upper boundary of the phase space. The addition of pile-up pushes jets out of the small $e_2^{(\beta)}$ region of phase space, where $D_2^{(\beta)}$ takes large values. Therefore, an effect of pile-up is to reduce the value the $D_2^{(\beta)}$ measured on a jet, compressing the long tail of the perturbative $D_2^{(\beta)}$ distribution (exhibited in Fig. 3-6, for example) toward a central value.

We can predict the value of $D_2^{(\beta)}$ for background jets in the limit of infinite pile-up. In this limit, the jet has a single hard core of radiation surrounded by perfectly uniform pile-up radiation. If the energy fraction of each of the n pile-up particles is z_{pu}, then the two- and three-point correlation functions take the values

$$e_2^{(\beta)} = n z_{pu} \,, \quad \text{and} \quad e_3^{(\beta)} = \binom{n}{2} z_{pu}^2 \,, \tag{3.43}$$

94

so that, as $n \to \infty$, we have the relation

$$e_3^{(\beta)} = \frac{1}{2}(e_2^{(\beta)})^2 \,. \tag{3.44}$$

Using the definition of $D_2^{(\beta)}$, we find that in this limit,

$$D_2^{(\beta)} = \frac{1}{2e_2^{(\beta)}} \,. \tag{3.45}$$

As the amount of pile-up increases, we expect that the distribution of $D_2^{(\beta)}$ accumulates about this value, with a minimal change in the mean, but significant decrease in the width of the distribution. This behavior is relatively distinct from that of most event shapes under pile-up, which tend to have a shift of the mean as pile-up is increased. The reason that this behavior is pronounced with $D_2^{(\beta)}$ is because it is both infrared and collinear unsafe and because the scalings of $e_2^{(\beta)}$ and $e_3^{(\beta)}$ in the observable and the upper boundary of the phase space are different.

For the background distribution of $C_2^{(\beta)}$, on the other hand, the parametric scaling of the observable is unaffected by the addition of pile-up, but we expect an $\mathcal{O}(1)$ shift of the mean of the distribution to larger values. As pile-up increases, from Eq. (3.44) we expect the distribution should accumulate about the infinite pile-up limit of $C_2^{(\beta)} = 1/2$. We therefore predict that as the pile-up increases, the distribution of $C_2^{(\beta)}$ on background jets becomes independent of β.

We can also understand the behavior of the signal distribution under the addition of soft pile-up radiation. As was discussed, and is shown schematically in Fig. 3-11, signal jets at larger $e_2^{(\beta)}$ are shifted vertically in the $(e_2^{(\beta)}, e_3^{(\beta)})$ phase space by pile-up radiation. This predicts that for both $C_2^{(\beta)}$ and $D_2^{(\beta)}$, the primary effect of the addition of pile-up radiation will be to shift the mean of the distribution to larger values, with a limited modification to its shape. Furthermore, due to the cubic contours for $D_2^{(\beta)}$, the shift of the mean of the distribution will be smaller for $D_2^{(\beta)}$ than for $C_2^{(\beta)}$, implying reduced sensitivity of $D_2^{(\beta)}$ to pile-up radiation.

Note that a similar analysis can be straightforwardly applied to the N-subjettiness observables $\tau_1^{(\beta)}$ and $\tau_2^{(\beta)}$. Because the analysis proceeds identically, we simply state the result. Under the addition of pile-up radiation, single prong jets at small $\tau_1^{(\beta)}$ experience a shift of both observables, but their parametric scaling remains the same:

$$\tau_1^{(\beta)} \to \tau_1^{(\beta)} + z_{pu} \, , \tag{3.46}$$

$$\tau_2^{(\beta)} \to \tau_2^{(\beta)} + z_{pu} \, . \tag{3.47}$$

That is, under the addition of pile-up, background jets move along the upper boundary of the phase space, where $\tau_2^{(\beta)} \sim \tau_1^{(\beta)}$.

For jets with two hard subjets, the value of $\tau_1^{(\beta)}$ is not affected, while $\tau_2^{(\beta)}$ shifts as

$$\tau_1^{(\beta)} \to \tau_1^{(\beta)} \, , \tag{3.48}$$

$$\tau_2^{(\beta)} \to \tau_2^{(\beta)} + z_{pu} \, . \tag{3.49}$$

This corresponds to a vertical movement in the $\tau_1^{(\beta)}, \tau_2^{(\beta)}$ phase space under the addition of pile-up, as was the case for $e_2^{(\beta)}, e_3^{(\beta)}$. We therefore expect a similar behavior for $\tau_{2,1}^{(\beta)}$ with the addition of pile-up, with a shift of the mean value for the signal distributions and accumulation near 1 for the background distributions. Unfortunately, power counting alone does not allow us to compare the expected shifts in $\tau_{2,1}^{(\beta)}$ as compared with those in $C_2^{(\beta)}$ and $D_2^{(\beta)}$.

Summary of Power Counting Predictions

Here, we summarize the main predictions from our power counting analysis of the impact of pile-up radiation on the $C_2^{(\beta)}$ and $D_2^{(\beta)}$ distributions. We have:

- For background $D_2^{(\beta)}$ distributions the primary effect of the addition of pile-up radiation is to narrow the distribution; in particular, the long tail of the $D_2^{(\beta)}$ distribution is truncated because pile-up moves those jets in the 1-prong region of phase space out of the region of small e_2. The peak of the $D_2^{(\beta)}$ background distribution should be

relatively insensitive to the addition of pile-up, with $D_2^{(\beta)}$ accumulating around $D_2^{(\beta)} = 1/(2e_2^{(\beta)})$ in the limit that uniform pile-up dominates.

- For background $C_2^{(\beta)}$ distributions the primary effect of the addition of pile-up radiation is a shift of the peak to larger values by an $\mathcal{O}(1)$ amount proportional to the pile-up. The distribution also becomes compressed, and accumulates around $C_2^{(\beta)} = 1/2$ in the limit that uniform pile-up dominates.

- For signal, the primary effect of the addition of pile-up radiation is to translate the mean of the distribution. The displacement of the mean is expected to be smaller for $D_2^{(\beta)}$ than for $C_2^{(\beta)}$ because of the different scalings for contours of constant $C_2^{(\beta)}$ and $D_2^{(\beta)}$.

Monte Carlo Analysis

We now study these predictions in Monte Carlo using the PYTHIA 8 event samples described in Sec. 3.3.1.[17] Pile-up was simulated by adding N_{PV} minimum bias events at the 8 TeV LHC, generated with PYTHIA 8, to the $pp \to Zj$ and $pp \to ZZ$ samples. To demonstrate the resilience of the distributions to pile-up, we wish to add pile-up radiation to a set of jets with well-defined perturbative properties. To do this, we cluster jets with the WTA recombination scheme [33, 89] and require that the mass of the jets in the absence of pile-up is $m_J < 100$ GeV. It was shown in Ref. [89] that the jet axis found by the WTA recombination scheme is robust to pile-up and so, when pile-up is included, the perturbative content of the jets will be unaffected. This procedure, although clearly not related to an experimental analysis, provides a measure of the sensitivity of the distributions to soft pile-up radiation. This procedure is similar to that used in Ref. [106] to assess the impact of pile-up and pile-up subtraction techniques on a variety of different jet shapes.

Using this sample, we can assess the degree to which the power counting predictions of Sec. 3.3.2 are realized in the Monte Carlo simulation. We begin by considering the effect of

---

[17]In this section, we restrict to the PYTHIA 8 generator, having satisfied ourselves in Sec. 3.3.1 that the parametrics of the perturbative phase space are well described by both generators.
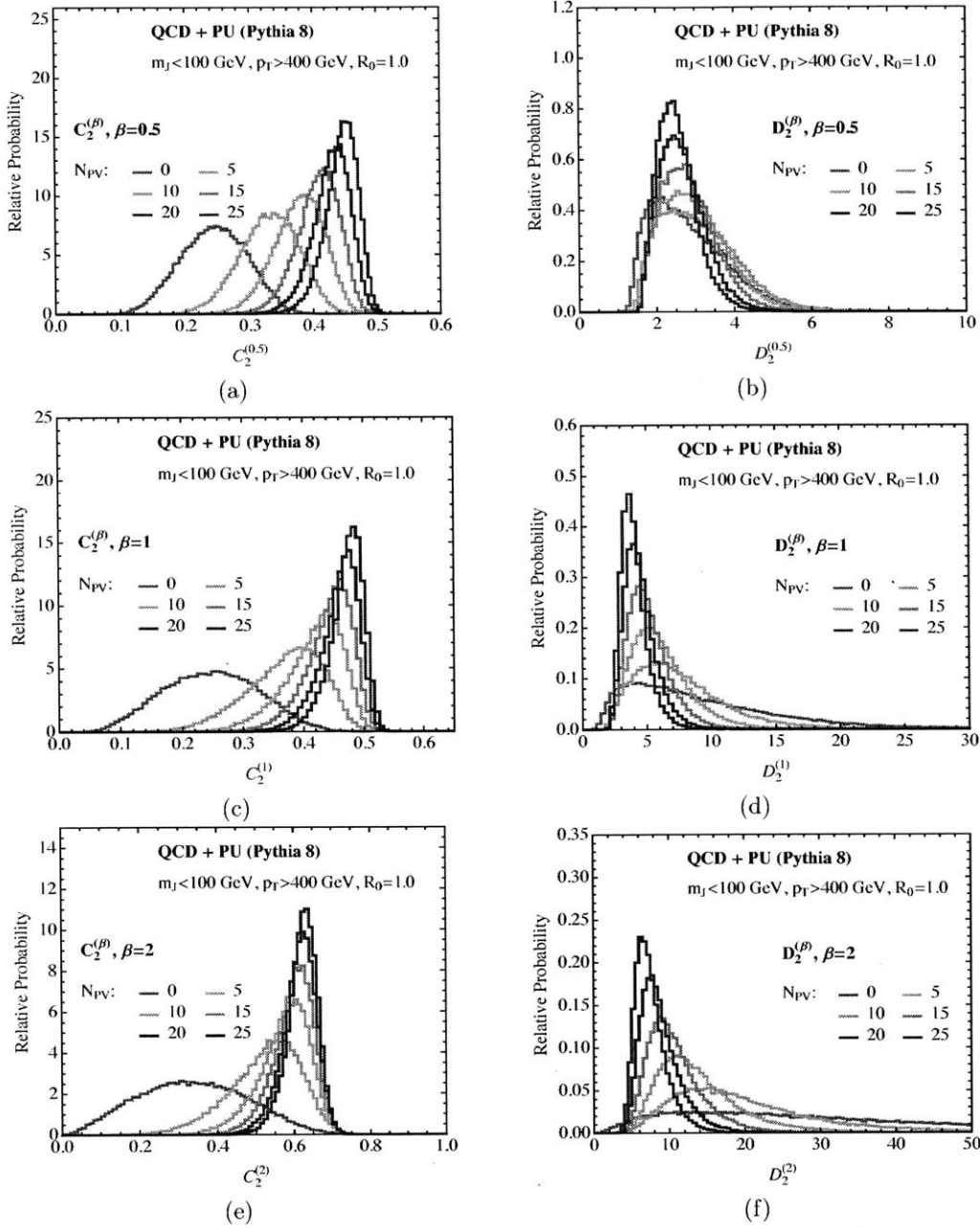
Figure 3-12: Effect of pile-up on the distributions of $C_2^{(\beta)}$ (left) and $D_2^{(\beta)}$ (right) for QCD jets for $\beta = 0.5, 1, 2$ as measured on the PYTHIA 8 samples. The number of pile-up vertices ranges from $N_{PV} = 0$ (no pile-up) to $N_{PV} = 25$.
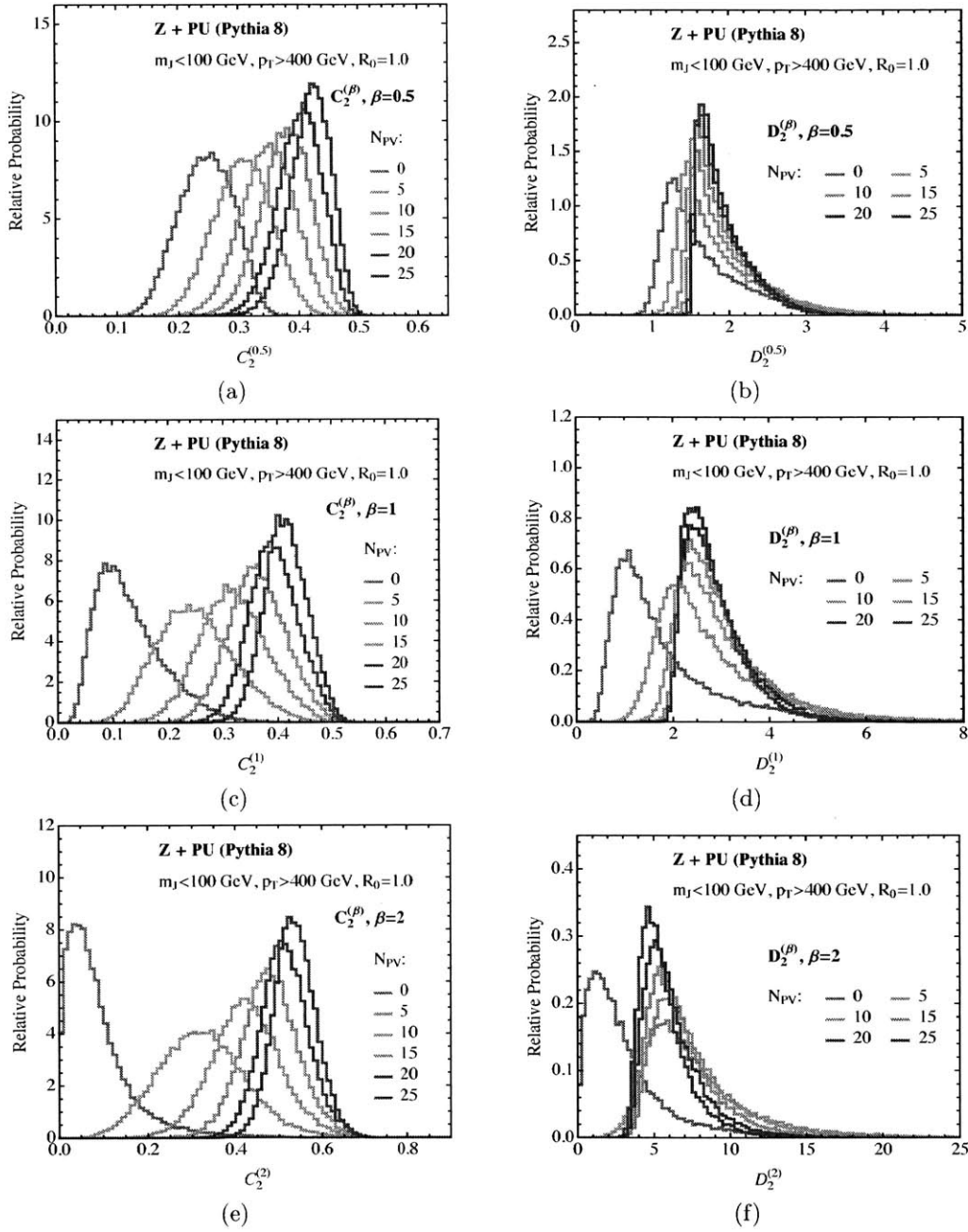
Figure 3-13: The same as Fig. 3-12, but measured on boosted Z jets.
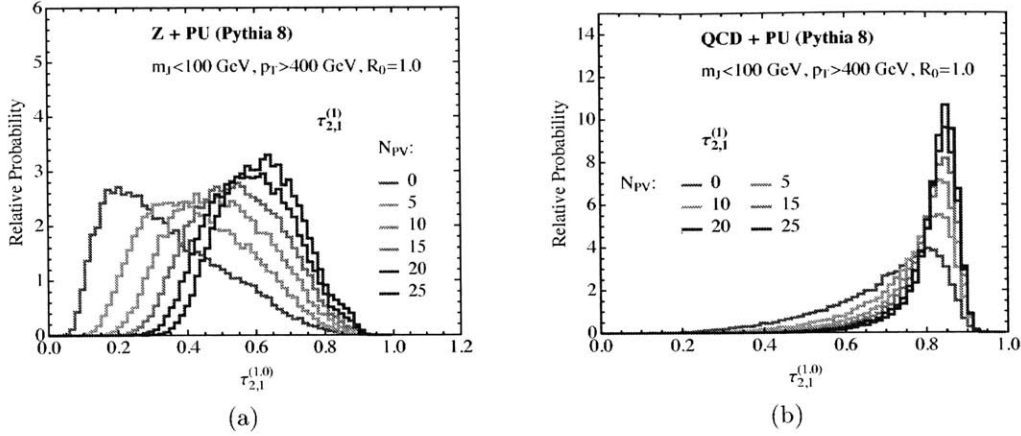
Figure 3-14: Effect of pile-up contamination on the measured value of $\tau_{2,1}^{(1)}$ for signal (left) and background (right) jets from the PYTHIA 8 samples. The number of pile-up vertices ranges from $N_{PV} = 0$ (no pile-up) to $N_{PV} = 25$.

pile-up on the background distributions. In Fig. 3-12, we plot background distributions for $C_2^{(\beta)}$ and $D_2^{(\beta)}$ with the addition of up to $N_{PV} = 25$ pile-up vertices for a few values of β. For the variable $D_2^{(\beta)}$, the power counting analysis of Sec. 3.3.2 predicted that the dominant effect of the addition of pile-up would be a compression of the long tail of the distribution into a peak around $1/2e_2^{(\beta)}$, with relatively little shift in the mean. This behavior is manifest for all three values of β shown. The peak value of the distribution is remarkably stable under the addition of pile-up. On the other hand, for the observable $C_2^{(\beta)}$, the mean of the distribution is highly unstable to the addition of pile-up. The dominant effects of pile-up on the distribution of $C_2^{(\beta)}$ is a displacement of the mean and the accumulation near the value $C_2^{(\beta)} = 1/2$.

In Fig. 3-13 we consider the same set of distributions as for Fig. 3-12, but for the signal boosted Z boson sample. In this case, the analysis of the phase space predicted that the dominant effect of the pile-up on the distributions is a shift for both $C_2^{(\beta)}$ and $D_2^{(\beta)}$, with the shift being smaller for $D_2^{(\beta)}$. This behavior is manifest in Fig. 3-13. The difference in the stability of the mean between $C_2^{(\beta)}$ and $D_2^{(\beta)}$ is particularly pronounced at small β. At larger β, the $D_2^{(\beta)}$ distribution exhibits a jump at small amounts of pile-up, and then remains

100

stable as pile-up increases. Unfortunately, we have not been able to understand this behavior completely from power counting. Nevertheless, the improved stability of the distributions of $D_2^{(\beta)}$ as compared with $C_2^{(\beta)}$ is promising.

For comparison, in Fig. 3-14, we consider the impact of pile-up on signal and background distributions for $\tau_{2,1}^{(\beta)}$, for the representative value $\beta = 1$. As for $C_2^{(\beta)}$ and $D_2^{(\beta)}$, we expect that the dominant effect of pile-up on the signal distributions is a shift of the peak value, while for the background distributions, we expect a small shift of the mean and an accumulation of the distribution near $\tau_{2,1}^{(\beta)} = 1$. This is exhibited in the Monte Carlo.

For a more quantitative study of the stability of the distributions to pile-up, we define

$$\delta_{X_2^{(\beta)}}(N_{PV}) = \frac{\langle X_2^{(\beta)}(N_{PV})\rangle - \langle X_2^{(\beta)}(N_{PV}=0)\rangle}{\sigma_{X_2^{(\beta)}}(N_{PV}=0)}, \qquad (3.50)$$

where $X_2^{(\beta)}$ stands for either $C_2^{(\beta)}$ or $D_2^{(\beta)}$ and σ denotes the standard deviation. This quantity is a measure of how much the mean of the distribution is affected by pile-up, normalized by the width of the distribution, which is important since the observables $C_2^{(\beta)}$ and $D_2^{(\beta)}$ have support over very different ranges. While it is clear from Fig. 3-12 that the dominant effect of pile-up on the background distributions for $D_2^{(\beta)}$ is not a shift of the mean, and so the change of the distribution is not accurately captured by the measure of Eq. (3.50), the deviation of the mean is a commonly studied measure of an observable's susceptibility to pile-up. In Fig. 3-15 we plot $\delta(N_{PV})$ for the variables $C_2^{(\beta)}$ and $D_2^{(\beta)}$. As was demonstrated in Figs. 3-12 and 3-13, the mean of the distributions of $D_2^{(\beta)}$ is considerably more stable for both the signal and background distributions.

3.4 Power Counting Quark vs. Gluon Discrimination

Unlike the case of boosted Z bosons vs. massive QCD jets, applying a power counting analysis to quark vs. gluon jet discrimination demonstrates the limitations of the technique. Both quark and gluon jets dominantly have only a single hard core, and so the natural

101
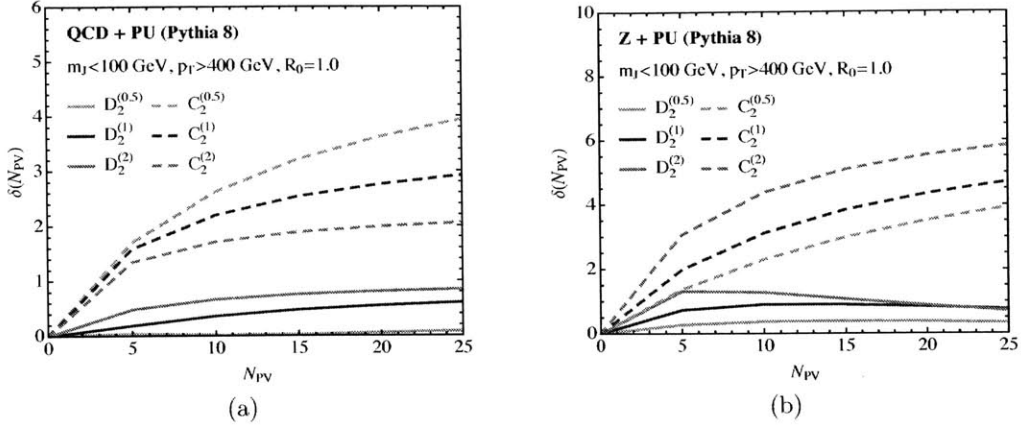
Figure 3-15: A comparison of the susceptibility as a function of the number of pile-up vertices N_{PV} of background (left) and signal (right) distributions for $C_2^{(\beta)}$ and $D_2^{(\beta)}$ to pile-up using the measure $\delta(N_{PV})$ for $\beta = 0.5, 1, 2$.

discrimination observables are the two-point energy correlation functions, $e_2^{(\beta)}$.[18] As shown in Sec. 3.2.2, power counting the two-point energy correlation functions constrains the soft and collinear radiation as:

$$e_2^{(\beta)} \sim z_s \sim R_{cc}^{\beta} . \tag{3.51}$$

With power counting alone, this is as far as our analysis can go. $e_2^{(\beta)}$ does not parametrically separate quark and gluon jets from one another.

This result is not surprising, however, because there are no qualities of quark and gluon jets that are parametrically different. Indeed,

$$C_A \sim C_F ,$$

$$N_C \sim n_f ,$$

$$\text{spin } 1 \sim \text{spin } 1/2 ,$$

where C_A and C_F are the color factors for gluons and quarks, N_C is the number of colors, and n_f the number of active fermions. Predictions of what the best observable for quark

---

[18]To next-to-leading logarithmic accuracy, $e_2^{(\beta)}$ are identical to the recoil-free angularities [33].

vs. gluon discrimination is requires a detailed analysis of the effects of these order-1 parameters, which has been studied in several chapters [122, 123, 30, 37]. However, with the additional input of the form of the splitting functions for quarks and gluons, we can predict that the discrimination power of $e_2^{(\beta)}$ improves as β decreases because smaller β emphasizes the collinear region of phase space over soft emissions. Collinear emissions are sensitive to the spin of the parton in addition to the total color of the jet, and thus are more distinct between quark and gluon jets. This prediction is borne out by explicit calculation to next-to-leading logarithmic accuracy [30]. An analytic calculation of the improved discrimination power from simultaneous measurement of the recoil-free angularities for two different powers of the angular exponent was calculated in [37].

Nevertheless, this suggests that power counting does make a definite prediction of quark vs. gluon discrimination performance. Because all the physics of quark vs. gluon jet discrimination is controlled by order-1 numbers, the predicted discrimination should be sensitive to the tuning of order-1 numbers in a Monte Carlo. It has been observed that PYTHIA 8 and HERWIG++ give wildly different predictions for quark vs. gluon discrimination power [30, 37, 64, 68], and presumably the difference is dominated by the tuning of the Monte Carlos. However, isolating pure samples of quark and gluon jets is challenging experimentally [64, 65, 68, 124] and most of the subtle differences between quarks and gluons only appear at an order formally beyond the accuracy of a Monte Carlo. Therefore, to solve this issue will require significant effort from experimentalists, Monte Carlo authors, and theorists to properly define quark and gluon jets, to identify the dominant physics, and to isolate pure samples for tuning.

3.5 Conclusions

In this chapter we have demonstrated that power counting techniques can be a powerful guiding principle when constructing observables for jet substructure and for understanding their behavior. Since power counting captures the parametric physics of the underlying theory, its predictions should be robust to Monte Carlo tunings. Using the simple example of

discriminating boosted Z bosons from QCD jets with the energy correlation functions, we showed that a power counting analysis identified $D_2^{(\beta)}$ as the natural discrimination observable. The scaling of this observable parametrically separates regions of the $(e_2^{(\beta)}, e_3^{(\beta)})$ phase space dominated by 1- and 2-prong jets. The distinction between 1- and 2-prong jets is invariant to boosts along the jet direction.

To verify the power counting predictions, we performed a Monte Carlo analysis comparing $D_2^{(\beta)}$ with a previously proposed observable, $C_2^{(\beta)}$, also formed from the energy correlation functions. We showed that $D_2^{(\beta)}$ is a superior observable for discrimination because $C_2^{(\beta)}$ inextricably mixes signal-rich and background-rich regions of phase space. All power counting predictions were confirmed by both HERWIG++ and PYTHIA 8, showing that the dominant behavior of the observables is governed by parametric scalings and not by $\mathcal{O}(1)$ numbers. This was contrasted with the case of quark vs. gluon discrimination for which no parametric differences exist, leading to large discrepancies when simulating quark vs. gluon discrimination with different Monte Carlo generators.

We also demonstrated that power counting can be used to understand the impact of pile-up on different regions of the phase space, and hence on the distributions of discriminating variables. The distributions for $D_2^{(\beta)}$ exhibited improved stability compared with those of $C_2^{(\beta)}$, while the background distributions have the interesting feature of being compressed to a central value by the addition of pile-up radiation.

We anticipate many directions to which the power counting approach could be applied. We have restricted ourselves in this chapter to a study of observables formed from ratios of energy correlation functions with the same angular exponent. A natural generalization is to ratios of energy correlation functions with different angular exponents, where the optimal observable is given by $D_2^{(\alpha,\beta)} = e_3^{(\beta)}/(e_2^{(\alpha)})^{3\beta/\alpha}$. Such variables could be useful when considering pile-up in the presence of mass cuts, which are required experimentally. In the presence of a mass cut, an angular exponent of e_2 near 2 provides a simple restriction on the phase space, while lowering the angular exponent of e_3 reduces the effect of soft wide angle radiation. Along these lines, the impact of grooming techniques on the phase space is also

simple to understand by power counting [90], and could be used to motivate the design of variables with desirable behavior under grooming.

As another example of considerable interest, the power counting analysis can be extended to the study of top quark discrimination variables by considering the phase space for 1-, 2- and 3-prong jets defined by the two-, three- and four-point energy correlation functions. While a complete analytic calculation for this case is not feasible, a power counting analysis is, and can be used to predict discriminating observables with considerably improved performance compared to those originally proposed in [30]. In the case of a three dimensional phase space, a cut on the jet mass only reduces the phase space to a two dimensional subspace, so that the functional form of the observable remains important. This case was studied in detail in [8], where an observable D_3 was proposed, and studied in Monte Carlo, where it showed improved performance over several other top tagging observables.

Our observation that boost-invariant combinations of the energy correlation functions are the most powerful discriminants can also be exploited for discrimination: we can use boost invariance as a guide for defining the best observables. Together with power counting, this gives a simple but powerful analytic handle to understand and design jet substructure observables.

Chapter 4

Analytic Boosted Boson Discrimination

Observables which discriminate boosted topologies from massive QCD jets are of great importance for the success of the jet substructure program at the Large Hadron Collider. Such observables, while both widely and successfully used, have been studied almost exclusively with Monte Carlo simulations. In this chapter we present the first all-orders factorization theorem for a two-prong discriminant based on a jet shape variable, D_2, valid for both signal and background jets. Our factorization theorem simultaneously describes the production of both collinear and soft subjets, and we introduce a novel zero-bin procedure to correctly describe the transition region between these limits. By proving an all orders factorization theorem, we enable a systematically improvable description, and allow for precision comparisons between data, Monte Carlo, and first principles QCD calculations for jet substructure observables. Using our factorization theorem, we present numerical results for the discrimination of a boosted Z boson from massive QCD background jets. We compare our results with Monte Carlo predictions which allows for a detailed understanding of the extent to which these generators accurately describe the formation of two-prong QCD jets, and informs their usage in substructure analyses. Our calculation also provides considerable insight into the discrimination power and calculability of jet substructure observables in general.

107

4.1 Introduction

The last several years has seen a surge of interest in the field of jet substructure [46, 47, 48, 125], both as an essential tool for extending new physics searches at the Large Hadron Collider (LHC) into the TeV energy regime, and as an important playground for improving our understanding of high energy QCD, both perturbative and non-perturbative. Of particular phenomenological interest are substructure observables that are sensitive to hard subjets within a jet. In the highly boosted regime, the hadronic decay products of electroweak-scale particles can become collimated and each appear as a jet in the detector. Unlike typical massive QCD jets, however, these boosted electroweak jets exhibit a multi-prong substructure that can be identified by the measurement of appropriate observables. Many such observables have been proposed and studied on LHC simulation or data [49, 50, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 126, 127] or used in new physics searches [71, 72, 73, 74, 75, 76, 77, 78, 79, 128, 129, 130, 131].

The vast majority of proposed jet substructure observables, however, have been analyzed exclusively within Monte Carlo simulation. While Monte Carlos play an essential role in the simulation of realistic hadron collision events, they can often obscure the underlying physics that governs the behavior of a particular observable. Additionally, it is challenging to disentangle perturbative physics from the tuning of non-perturbative physics so as to understand how to systematically improve the accuracy of the Monte Carlo. Recently, there has been an increasing number of analytical studies of jet substructure observables, including the calculation of the signal distribution for N-subjettiness to next-to-next-to-next-to-leading-log order [85], a fixed-order prediction for planar flow [132], calculations of groomed jet masses [86, 87, 37, 133] and the jet profile/ shape [134, 135, 136, 137, 138, 139, 140] for both signal and background jets, an analytic understanding of jet charge [141, 142], predictions for fractional jet multiplicity [143], and calculations of the associated subjet rate [144]. Especially in the case of the groomed jet observables, analytic predictions informed the construction of more performant and easier to calculate observables. With the recent start of Run 2 of the LHC, where the phase space for high energy jets only grows, it will be increasingly important

108

to have analytical calculations to guide experimental understanding of jet dynamics.

It is well known that the measurement of observables on a jet can introduce ratios of hierarchical scales appearing in logarithms at every order in the perturbative expansion. Accurate predictions over all of phase space require resummation of these large logarithms to all orders in perturbation theory. While this resummation is well understood for simple observables such as the jet mass [80, 81, 82, 83, 84], where it has been performed to high accuracy, a similar level of analytic understanding has not yet been achieved for more complicated jet substructure observables. Jet substructure observables are typically sensitive to a multitude of scales, corresponding to characteristic features of the jet, resulting in a much more subtle procedure for resummation.

A ubiquitous feature of some of the most powerful observables used for identification of jet substructure is that they are formed from the ratio of infrared and collinear (IRC) safe observables. Examples of such observables include ratios of N-subjettiness variables [32, 31], ratios of energy correlation functions [30, 5, 8], or planar flow [145]. In general, ratios of IRC safe observables are not themselves IRC safe [106] and cannot be calculated to any fixed order in perturbative QCD. Nevertheless, it has been shown that these ratio observables are calculable in resummed perturbation theory and are therefore referred to as Sudakov safe [88, 40, 89, 146]. Distributions of Sudakov safe observables can be calculated by appropriately marginalizing resummed multi-differential cross sections of IRC safe observables. An understanding of the factorization properties of multi-differential jet cross sections has been presented in Refs. [10, 34, 11] by identifying distinct factorization theorems in parametrically separated phase space regions defined by the measurements performed on the jet. Combining this understanding of multi-differential factorization with the required effective field theories, all ingredients are now available for analytic resummation and systematically improvable predictions.

As an explicit example, observables that resolve two-prong substructure are sensitive to both the scales characterizing the subjets as well as to the scales characterizing the full jet. A study of the resummation necessary for describing jets with a two-prong substructure

was initiated in Ref. [105] which considered the region of phase space with two collinear subjets of comparable energy, and introduced an effective field theory description capturing all relevant scales of the problem. Recently, an effective field theory description for the region of two-prong jet phase space with a hard core and a soft, wide angle subjet was developed in Ref. [11], where it was applied to the resummation of non-global logarithms [38]. Combined, the collinear subjet and soft subjet factorization theorems allow for a complete description of the dominant dynamics of jets with two-prong substructure.

In this chapter we will study the factorization and resummation of the jet substructure observable D_2 [5], a ratio-type observable formed from the energy correlation functions. We will give a detailed effective theory analysis using the language of soft-collinear effective theory (SCET) [22, 23, 24, 25] in all regions of phase space required for the description of a one or two-prong jet, and will prove all-orders leading-power factorization theorems in each region. We will then use these factorization theorems to calculate the D_2 distribution for jets initiated by boosted hadronic decays of electroweak bosons or from light QCD partons and compare to Monte Carlo simulation. These calculations will also allow us to make first-principles predictions for the efficiency of the observable D_2 to discriminate boosted electroweak signal jets from QCD background jets.

Our factorized description is valid to all orders in α_s, expressing the cross section as a product of field theoretic matrix elements, each of which is calculable order by order in perturbation theory, allowing for a systematically improvable description of the D_2 observable. Furthermore, the factorization theorem enables a clean separation of perturbative and non-perturbative physics, allowing for non-perturbative contributions to the observable to be included in the analytic calculation through the use of shape functions [147, 148]. In this chapter we work to next-to-leading logarithmic (NLL) accuracy to demonstrate all aspects of the required factorization theorems necessary for precision jet substructure predictions. We will see that even at this first non-trivial order, we gain insight into qualitative and quantitative features of the D_2 distribution. While we will give an extensive discussion of our numerical results and comparisons with a variety of Monte Carlo programs in this
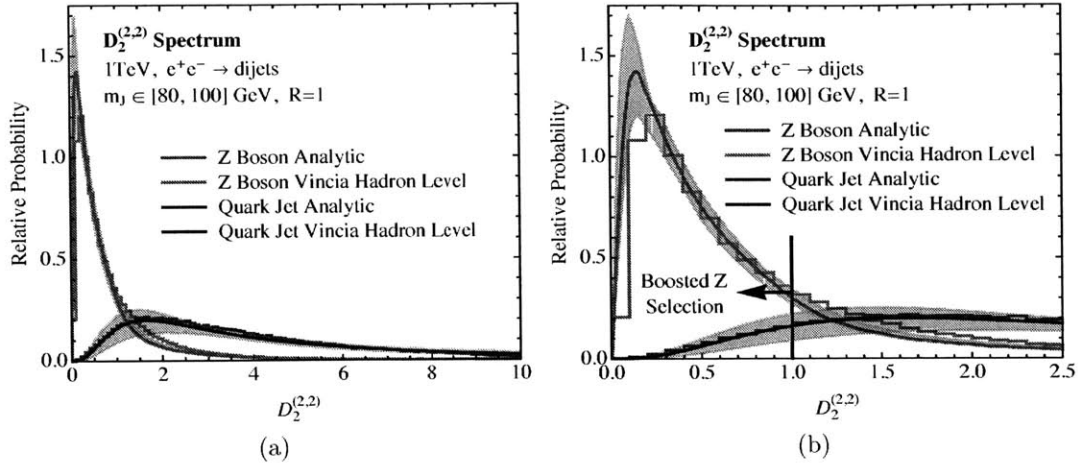
Figure 4-1: Comparison of our analytic calculation with VINCIA Monte Carlo predictions for the two prong discriminant, D_2. Predictions for both boosted Z bosons and massive QCD jets at a 1 TeV e^+e^- collider are shown. The Monte Carlo is fully hadronized, and non-perturbative effects have been included in the analytic calculation through a shape function. In a) we show the complete distribution, and in b) we zoom in to focus on the region relevant for boosted Z discrimination.

chapter, in Fig. 4-1 we compare our analytic predictions for the D_2 observable, including non-perturbative effects, for hadronically-decaying boosted Z bosons and QCD jets in e^+e^- collisions with the distributions predicted by the VINCIA [149, 150, 151, 152, 153, 154] Monte Carlo program at hadron level. Excellent agreement between analytic and Monte Carlo predictions is observed, demonstrating a quantitative understanding of boosted jet observables from first principles.

4.1.1 Overview of the chapter

While there exists a large number of two-prong discriminants in the jet substructure literature, any of which would be interesting to understand analytically, we will use calculability and factorizability as guides for constructing the observable to study in this chapter. This procedure will ultimately lead us to the observable D_2 and will demonstrate that D_2 has particularly nice factorization and calculability properties. This approach will proceed in the following steps:

111

1. Identify the relevant subjet configurations for the description of a two-prong discriminant.

2. Isolate each of these relevant regions by the measurement of a collection of IRC safe observables.

3. Study the phase space defined by this collection of IRC safe observables, and prove all-orders factorization theorems in each parametrically-defined region of phase space.

4. Identify a two-prong discriminant formed from the collection of IRC safe observables which respects the parametric factorization theorems of the phase space.

A detailed analysis of each of these steps will be the subject of this chapter. Here, we provide a brief summary so that the logic of the approach is clear, and so that the reader can skip technical details in the different sections without missing the general idea of the approach.

The complete description of an observable capable of discriminating one- from two-prong substructure requires the factorized description of the following three relevant subjet configurations, shown schematically in Fig. 4-2:

- **Soft Haze**: Fig. 4-2a shows a jet in what we will refer to as the soft haze region of phase space. In the soft haze region there is no resolved subjet, only a single hard core with soft wide angle emissions. This region of phase space typically contains emissions beyond the strongly ordered limit, but is the dominant background region for QCD jets, for which a hard splitting is α_s suppressed.

- **Collinear Subjets**: Fig. 4-2b shows a jet with two hard, collinear subjets. Both subjets carry approximately half of the total energy of the jet, and have a small opening angle. This region of phase space, and its corresponding effective field theory description, has been studied in Ref. [105].

- **Soft Subjet**: Fig. 4-2c shows the soft subjet region of phase space which consists of jets with two subjets with hierarchical energies separated by an angle comparable to the jet radius R. The soft subjet probes the boundary of the jet and we take $R \sim 1$.

112
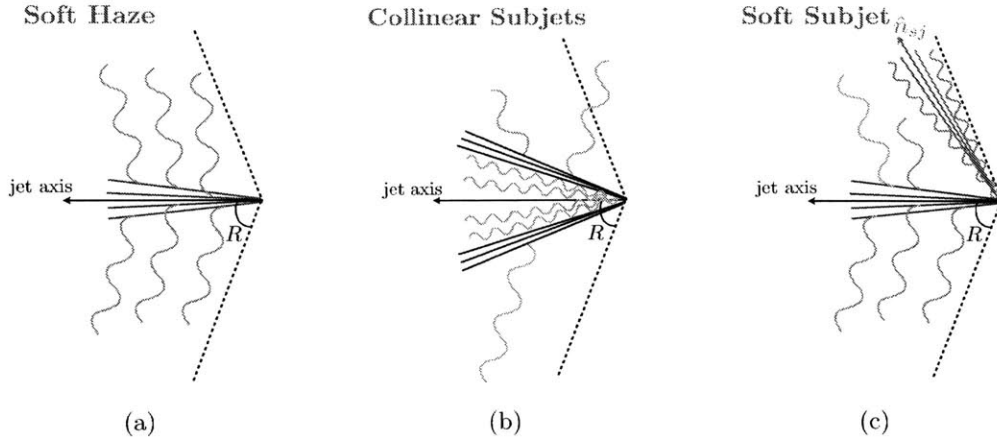
Figure 4-2: Regions of interest for studying the two-prong substructure of a jet. a) Soft haze region in which no subjets are resolved. b) Collinear subjets with comparable energy and a small opening angle. c) Soft subjet carrying a small fraction of the total energy, and at a wide angle from the hard subjet.

An effective field theory description for this region of phase space was presented in Ref. [11].

As a basis of IRC safe observables for isolating these three subjet configurations, we use the energy correlation functions [30], which we define in Sec. 4.2.1. In particular, we will show that the measurement of three energy correlation functions, two 2-point energy correlation functions $e_2^{(\alpha)}, e_2^{(\beta)}$, and a 3-point energy correlation function $e_3^{(\alpha)}$, allows for parametric separation of the different subjet configurations. While we will focus on the particular case of observables formed from the energy correlation functions, we believe that this approach is more general and could be applied to other IRC safe observable bases.

With the energy correlation functions as our basis, we study the multi-differential phase space defined by the simultaneous measurement of these observables on a jet in Sec. 4.2. Using the power counting technique of Refs. [5, 8], we show that the angular exponents of the energy correlation functions, α and β, can be chosen such that the different subjet configurations occupy parametrically separated regions of this phase space, and extend to all boundaries of the phase space. This parametric separation allows for each region to be

separately described by its own effective field theory. The required effective field theories are described in Sec. 4.3, and are formulated in the language of SCET. The formulation in SCET allows us to prove all-orders factorization theorems valid at leading-power in each of the phase space regions, and to resum logarithms to arbitrary accuracy using renormalization group techniques.

Having understood in detail both the structure of the phase space defined by the IRC safe measurements as well as the factorization theorems defined in each region, we will show in Sec. 4.4.1 that this leads unambiguously to the definition of a two-prong discriminant observable which is amenable to factorization. This observable will be a generalized form of D_2 [5] which will depend on both angular exponents α and β. Calculating the distribution of D_2 is accomplished by appropriate marginalization of the multi-differential cross section. Depending on the phase space cuts that have been made, D_2 may or may not be IRC safe itself, and so the marginalization will in general only be defined within resummed perturbation theory.

The outline of this chapter is as follows. In Sec. 4.2 we define the energy correlation functions used in this chapter and describe how the different subjet configurations shown schematically in Fig. 4-2 can be isolated by demanding parametric relations between the measured values of these observables. In Sec. 4.3 we discuss the effective field theory descriptions in the different phase space regions, and present the factorization theorems that describe their dynamics. Although some of the relevant effective field theories have been presented elsewhere, we attempt to keep the discussion self-contained by providing a brief review of their most salient features. All field theoretic definitions of the functions appearing in the factorization theorems, as well as their calculations to one-loop accuracy, are provided in appendices.

In Sec. 4.4 we show how the detailed understanding of the multi-differential phase space leads to the definition of D_2 as a powerful one- versus two-prong jet discriminant. In Sec. 4.4.2 we emphasize that without a mass cut, D_2 is not IRC safe but is Sudakov safe and whose all-orders distribution exhibits paradoxical dependence on α_s. In Sec. 4.4.3 we study the

114

fixed-order distribution of D_2 in the presence of a mass cut to understand its behavior in singular limits. In Sec. 4.4.4 we discuss how the different effective field theories can be consistently merged to give a factorized description of the D_2 observable, and introduce a novel zero-bin procedure to implement this merging.

In Sec. 4.5 we present numerical results for both signal and background distributions for D_2 as measured in e^+e^- collisions and compare our analytic calculation with several Monte Carlo generators. We emphasize many features of the calculation which provide considerable insight into two-prong discrimination, and the ability of current Monte Carlo generators to accurately describe substructure observables. In Sec. 4.6 we discuss numerical results for the D_2 observable at e^+e^- collisions at the Z pole at LEP, and demonstrate that being sensitive to correlations between three emissions, the D_2 observable can be used as a more differential probe of the perturbative shower for tuning Monte Carlo generators. In Sec. 4.7 we discuss how to extend our calculations to pp collisions at the LHC. We conclude in Sec. 4.8, and discuss future directions for further improving the analytic understanding of jet substructure.

4.2 Characterizing a Two-Prong Jet

In this section, we develop the framework necessary to construct the all-orders factorization theorems for analytic two-prong discrimination predictions. We begin in Sec. 4.2.1 by defining the energy correlation functions, which we will use to isolate the three subjet configurations discussed in the introduction. Using the power counting analysis of Refs. [5, 8], we study the phase space defined by measuring the energy correlation functions in Sec. 4.2.2. Throughout this chapter, our proxy for a two-prong jet will be a boosted, hadronically decaying Z boson, but our analysis holds for W or H bosons, as well.

4.2.1 Observable Definitions

To distinguish the three different subjet configurations of Fig. 4-2 with IRC safe measurements, observables which are sensitive to both one- and two-prong structure are required.

Although many possible observable bases exist, in this chapter we will use the energy correlation functions [30, 5], as we will find that they provide a convenient basis.

The n-point energy correlation function is an IRC safe observable that is sensitive to n-prong structure in a jet. For studying the two-prong structure of a jet, we will need the 2- and 3-point energy correlation functions, which we define for e^+e^- collisions as [30][1]

$$e_2^{(\alpha)} = \frac{1}{E_J^2} \sum_{i<j \in J} E_i E_j \left(\frac{2p_i \cdot p_j}{E_i E_j} \right)^{\alpha/2} ,$$

$$e_3^{(\alpha)} = \frac{1}{E_J^3} \sum_{i<j<k \in J} E_i E_j E_k \left(\frac{2p_i \cdot p_j}{E_i E_j} \frac{2p_i \cdot p_k}{E_i E_k} \frac{2p_j \cdot p_k}{E_j E_k} \right)^{\alpha/2} .$$

(4.1)

Here J denotes the jet, E_i and p_i are the energy and four momentum of particle i in the jet and α is an angular exponent that is required to be greater than 0 for IRC safety. The 4-point and higher energy correlation functions are defined as the natural generalizations of Eq. (4.1), although we will not use them in this chapter.

While we will mostly focus on the case of an e^+e^- collider, the energy correlation functions have natural generalizations to hadron colliders, by replacing E by p_T and using hadron collider coordinates, η and ϕ. This definition is given explicitly in Eq. (4.81). At central rapidity, this modification does not change the behavior of the observables, or any of the conclusions presented in the next sections. Of course, the hadron collider environment has other effects not present in an e^+e^- collider, like initial state radiation or underlying event, that will affect the energy correlation functions. A brief discussion of the behavior of the energy correlation functions in pp colliders will be given in Sec. 4.7. Numerical implementations of the energy correlation functions for both e^+e^- and hadron colliders are available in the EnergyCorrelator FASTJET CONTRIB [100, 101].

---

[1] For massive hadrons, there exist several possible definitions of the energy correlation functions depending on the particular mass scheme [155, 156]. The definition in Eq. (4.1) is an E-scheme definition. A p-scheme definition will be presented in Sec. 4.6 when we discuss the connection to LEP. Since the different definitions are equivalent for massless partons, their perturbative calculations are identical. The different definitions differ only in their non-perturbative corrections.

4.2.2 Power Counting the $e_2^{(\alpha)}, e_2^{(\beta)}, e_3^{(\alpha)}$ Phase Space

With a basis of IRC safe observables identified, we now demonstrate that the measurement of multiple energy correlation functions parametrically separates the three different subjet configurations identified in Fig. 4-2. In particular, the simultaneous measurement of $e_2^{(\alpha)}$, $e_2^{(\beta)}$, and $e_3^{(\alpha)}$ is sufficient for this purpose, and we will study in detail the phase space defined by their measurement. From this analysis, we will be able to determine for which values of the angular exponents α and β the three subjet configurations are parametrically separated within this phase space. The power counting parameters that define "parametric" will be set by the observables themselves, as is typical in effective field theory.

We begin by considering how the energy correlation functions can be used to separate one- and two-prong jets. This has been previously discussed in Ref. [5] by measuring $e_2^{(\alpha)}$ and $e_3^{(\alpha)}$, but here we consider the phase space defined by $e_2^{(\beta)}$ and $e_3^{(\alpha)}$ with α and β in general different. A minimal constraint on the angular exponents, both for calculability and discrimination power, is that the soft haze and collinear subjets configurations are parametrically separated by the measurements. A power counting analysis of the soft subjet region yields no new constraints beyond those from the soft haze or collinear subjets.

The setup for the power counting analysis of the soft haze and collinear subjets configurations is shown in Fig. 4-3, where all relevant modes are indicated. The one-prong jet illustrated in Fig. 4-3a is described by soft modes with energy fraction z_s emitted at $\mathcal{O}(1)$ angles, and collinear modes with characteristic angular size θ_{cc} with $\mathcal{O}(1)$ energy fraction. The collinear subjets configuration illustrated in Fig. 4-3b consists of two subjets, each of which carry an $\mathcal{O}(1)$ fraction of the jet's energy and are separated by an angle $\theta_{12} \ll 1$. Each of the subjets has collinear emissions at a characteristic angle $\theta_{cc} \ll \theta_{12}$, and global soft radiation at large angles with respect to the subjets, with characteristic energy fraction $z_s \ll 1$. In the case of two collinear subjets arising from the decay of a color singlet particle, the long wavelength global soft radiation is not present due to color coherence, but the power counting arguments of this section remain otherwise unchanged. Finally, there is radiation from the dipole formed from the two subjets (called "collinear-soft" radiation), with charac-

117
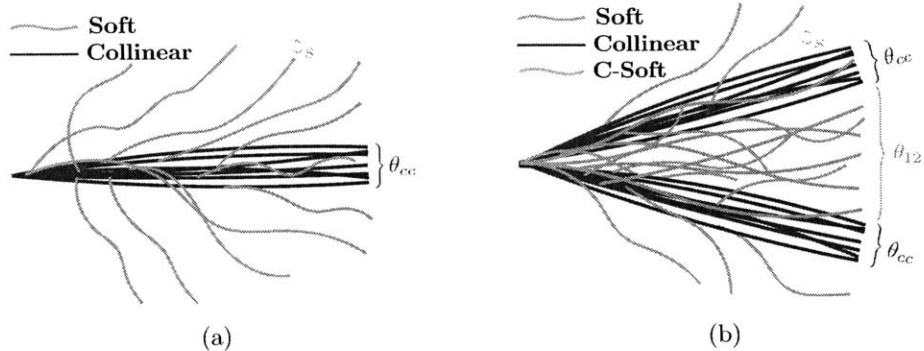
Figure 4-3: a) Schematic of a one-prong soft haze jet, dominated by collinear (blue) and soft (green) radiation. The angular size of the collinear radiation is θ_{cc} and the energy fraction of the soft radiation is z_s. b) Schematic of a jet resolved into two collinear subjets, dominated by collinear (blue), soft (green), and collinear-soft (orange) radiation emitted from the dipole formed by the two subjets. The subjets are separated by an angle θ_{12} and the energy fraction of the collinear-soft radiation is z_{cs}.

teristic angle θ_{12} from the subjets, and with energy fraction z_{cs}. The effective theory of this phase space region for the observable N-jettiness [104] was studied in Ref. [105].[2]

We are now able to determine the parametric form of the dominant contributions to the observables $e_2^{(\beta)}$ and $e_3^{(\alpha)}$. In the soft haze region, the dominant contributions to the energy correlation functions are[3]

$$
\begin{aligned}
e_2^{(\beta)} &\sim z_s + \theta_c^\beta \,, \\
e_3^{(\alpha)} &\sim \theta_c^{3\alpha} + z_s^2 + \theta_c^\alpha z_s \,.
\end{aligned}
\tag{4.2}
$$

---

[2]It is of historical interest to note that the generalization of two-prong event shapes, such as thrust, to event shapes for characterizing three jet structure was considered early on, for example with the introduction of the triplicity event shape [157]. However, it was not until more recently, with the growth of the jet substructure field at the LHC, that significant theoretical study was given to such observables.

[3]It is important to understand that this relationship is valid to an arbitrary number of emissions. When performing the power counting, a summation over all the particles with soft and collinear scalings in the jet must be considered. However, to determine the scalings of the observable, it is sufficient to consider the scaling of the different types of individual terms in the sum. For example, the three terms contributing to the expression for $e_3^{(\alpha)}$ arise from correlations between subsets of three collinear particles, one collinear particle and two soft particles, and two collinear particles and a soft particle, respectively. Contributions from other combinations of particles are power suppressed. Because of this simplification, in this chapter we will never write explicit summations when discussing the scaling of observables.

118

	$\alpha \leq \beta/2$	$\beta/2 < \alpha < 2\beta/3$	$2\beta/3 \leq \alpha < \beta$	$\alpha \geq \beta$
upper	$e_3^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^{3\alpha/\beta}$	$e_3^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^{\alpha/\beta+1}$	$e_3^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^{\alpha/\beta+1}$	$e_3^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^{2}$
lower	$e_3^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^{2}$	$e_3^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^{2}$	$e_3^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^{3\alpha/\beta}$	$e_3^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^{3\alpha/\beta}$

Table 4.1: Parametric scaling of the upper and lower boundaries of the one-prong region of the $e_2^{(\beta)}, e_3^{(\alpha)}$ phase space as a function of the angular exponents α and β.

From these parametrics, it is straightforward to show that one-prong jets live in a region of the $e_2^{(\beta)}, e_3^{(\alpha)}$ phase space bounded from above and below, whose precise scaling depends on the relative size of the angular exponents α and β. The scaling of upper and lower boundaries of the one-prong region of phase space for all α and β are listed in Table 4.1. For $\alpha = \beta$, as studied in Ref. [5], one-prong jets live in the region defined by $\left(e_2^{(\beta)}\right)^3 \lesssim e_3^{(\beta)} \lesssim \left(e_2^{(\beta)}\right)^2$.

For the collinear subjets configuration, the dominant contributions to the observables $e_2^{(\beta)}$ and $e_3^{(\alpha)}$ are

$$e_2^{(\beta)} \sim \theta_{12}^{\beta},$$
$$e_3^{(\alpha)} \sim \theta_{cc}^{\alpha}\theta_{12}^{2\alpha} + \theta_{12}^{\alpha}z_s + \theta_{12}^{3\alpha}z_{cs} + z_s^2. \tag{4.3}$$

The 2-point energy correlation function $e_2^{(\beta)}$ is set by the angle of the hard splitting, θ_{12}, and the scaling of all other modes (soft, collinear, or collinear-soft) are set by the $e_3^{(\alpha)}$ measurement. The requirement

$$z_{cs} \sim \frac{e_3^{(\alpha)}}{\left(e_2^{(\beta)}\right)^{3\alpha/\beta}} \ll 1, \tag{4.4}$$

then implies that the two-prong jets occupy the region of phase space defined by $e_3^{(\alpha)} \ll \left(e_2^{(\beta)}\right)^{3\alpha/\beta}$.

For optimal discrimination, the one- and two-prong regions of this phase space should not overlap. Since they are physically distinct, a proper division of the phase space will allow distinct factorizations, simplifying calculations. Comparing the boundaries of the one-prong

119

region listed in Table 4.1 with the upper boundary of the two-prong region from Eq. (4.4), we find that the one- and two-prong jets do not overlap with the following restriction on the angular exponents α and β:

$$3\alpha \geq 2\beta .$$ (4.5)

Note that when $\alpha = \beta$ this is satisfied, consistent with the analysis of Ref. [5].

Because these power counting arguments rely exclusively on the parametric behavior of QCD in the soft and collinear limits, they must be reproduced by any Monte Carlo simulation, regardless of its shower and hadronization models. To illustrate the robust boundary between the one- and two-prong regions of phase space predicted in Eq. (4.5), in Fig. 4-4, we plot the distribution in the $e_2^{(\beta)}, e_3^{(\alpha)}$ plane of jets initiated by light QCD partons and those from boosted hadronic decays of Z bosons as generated in e^+e^- collisions in PYTHIA [108, 109]. Details of the Monte Carlo generation are presented in Sec. 4.5. QCD jets are dominantly one-pronged, while jets from Z decays are dominantly two-pronged. We have chosen to use angular exponents $\alpha = \beta = 1$ for this plot, as the small value of the angular exponent allows the structure of the phase space to be seen in a non-logarithmic binning. The predicted behavior persists for all values of α and β consistent with Eq. (4.5), while the choice made here is simply for illustrative aesthetics. On these plots, we have added dashed lines corresponding to the predicted one- and two-prong phase space boundaries to guide the eye. The one-prong QCD jets and the two-prong boosted Z jets indeed dominantly live in their respective phase space regions as predicted by power counting.

The measurement of $e_2^{(\beta)}$ and $e_3^{(\alpha)}$ alone is sufficient to separate one- and two prong jets. However, the two-prong jets can exhibit either collinear subjets or a soft, wide angle subjet. To separate the collinear and soft subjet two-prong jets, we make an additional IRC safe measurement on the full jet. Following Ref. [11], in addition to $e_2^{(\beta)}$ and $e_3^{(\alpha)}$, we measure $e_2^{(\alpha)}$, with $\alpha \neq \beta$. In particular, the soft subjet and collinear subjet regions of phase space
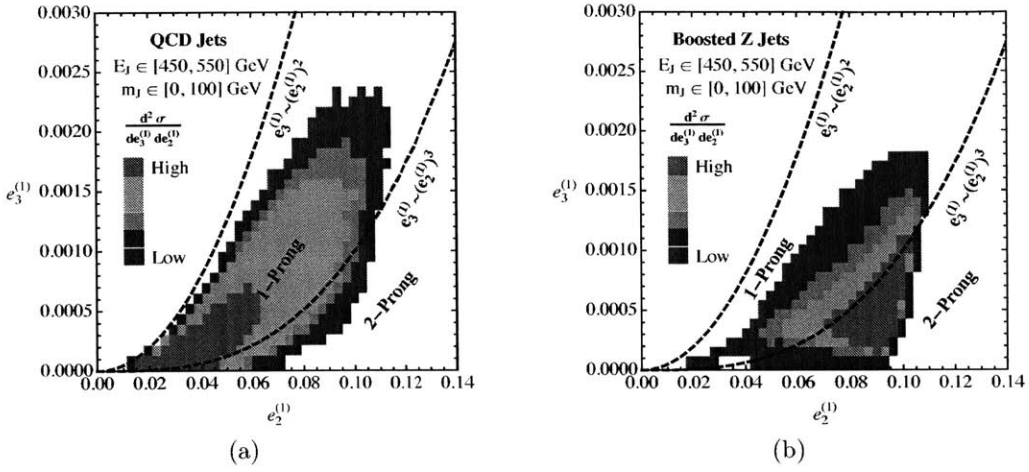
Figure 4-4: Monte Carlo distributions in the $e_2^{(1)}, e_3^{(1)}$ plane, for QCD quark jets (left) and boosted $Z \to q\bar{q}$ jets (right). The parametric scalings predicted by the power counting analysis are shown as dashed lines, and the one- and two-prong regions of phase space are labelled, and extend between the parametric boundaries. Note the upper boundary is constrained to have a maximal value of $\frac{1}{2}(e_2^{(\alpha)})^2 = e_3^{(\alpha)}$.

are defined by the simple conditions

$$\text{Collinear Subjet:} \qquad e_2^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^{\alpha/\beta}, \qquad (4.6)$$

$$\text{Soft Subjet:} \qquad e_2^{(\alpha)} \sim e_2^{(\beta)}. \qquad (4.7)$$

For $\alpha \neq \beta$ and $e_2^{(\beta)} \ll 1$, these two regions are parametrically separated. Equivalently, in the two-prong region of phase space the measurement of both $e_2^{(\alpha)}$ and $e_2^{(\beta)}$ can be used to give IRC safe definitions to the subjet energy fraction and splitting angle, allowing the soft subjet and collinear subjets to be distinguished. In Fig. 4-5 we summarize and illustrate the measurements that we make on the jet and the parametric relations between the measured values of the energy correlation functions that define the three phase space regions. The phase space plots of Figs. 4-5b and 4-5c were also presented in Ref. [11].

121

Subjet Configuration	Defining Relation
Soft Haze	$\left(e_2^{(\beta)}\right)^{3\alpha/\beta} \lesssim e_3^{(\alpha)} \lesssim \left(e_2^{(\beta)}\right)^2$
Collinear Subjets	$e_2^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^{\alpha/\beta}$ and $e_3^{(\alpha)} \ll \left(e_2^{(\beta)}\right)^{3\alpha/\beta}$
Soft Subjet	$e_2^{(\alpha)} \sim e_2^{(\beta)}$ and $e_3^{(\alpha)} \ll \left(e_2^{(\beta)}\right)^{3\alpha/\beta}$

(a)

(b)

(c)
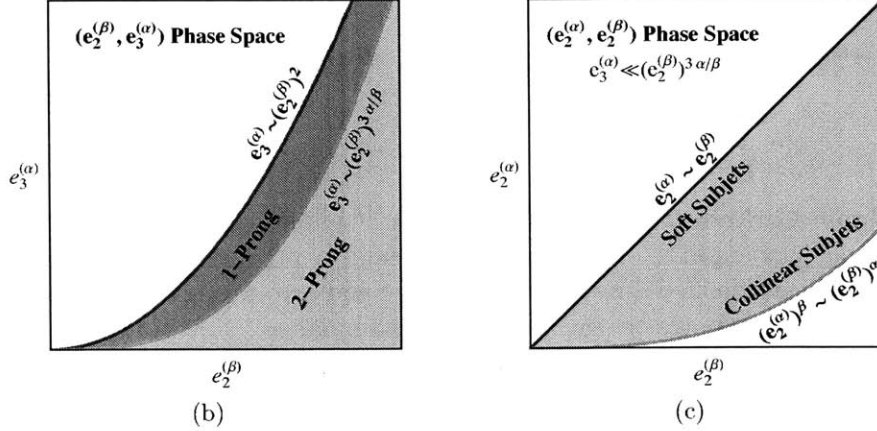
Figure 4-5: a) Table summarizing the defining relations for the different subjet configurations in terms of the energy correlation functions $e_2^{(\alpha)}, e_2^{(\beta)}, e_3^{(\alpha)}$. b) The one- and two-prong jets regions in the $e_2^{(\beta)}, e_3^{(\alpha)}$ phase space. Jets with a two-prong structure lie in the lower (orange) region of phase space, while jets with a one-prong structure lie in the upper (purple) region of phase space. c) The projection onto the $e_2^{(\alpha)}, e_2^{(\beta)}$ phase space in which the soft subjet and collinear subjets are separated.

Jet Mass Cuts

In addition to discriminating QCD jets from boosted Z bosons by their number of resolved prongs, we must also impose a mass cut on the jet to ensure that the jet is compatible with a Z decay. To include a mass cut in our analysis, for general angular exponents α and β, we would need to measure four observables on the jet: $e_2^{(\alpha)}$, $e_2^{(\beta)}$, $e_3^{(\alpha)}$ and the jet mass. This would significantly complicate calculations and introduce new parametric phase space regions that would need to be understood. To avoid this difficulty, we note that, for our

122

definition of $e_2^{(\alpha)}$ from Eq. (4.1), if all final state particles are massless, then

$$e_2^{(2)} = \frac{m_J^2}{E_J^2},\qquad(4.8)$$

where m_J is the mass of the jet. Therefore, choosing $\beta = 2$ we can trivially impose a mass cut within the framework developed here. Throughout the rest of this chapter, we will set $\beta = 2$ for this reason. Importantly, from Monte Carlo studies it has been shown that $\beta \sim 2$ provides optimal discrimination power [30, 5], so this restriction does not limit the phenomenological relevance of our results.

Substituting the value $\beta = 2$ into the power counting condition of Eq. (4.5), we find that the one- and two-prong regions of phase space are separated if

$$\alpha \geq \frac{4}{3}.\qquad(4.9)$$

To achieve a parametric separation of the one- and two-prong regions of phase space, we will demand that the scalings defining the different regions be separated by at least a single power of $e_2^{(\beta)}$. For example, choosing $\alpha = \beta = 2$, the scalings of the one-prong and two-prong regions are $e_3^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^3$ and $e_3^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^2$, which are parametrically different. We therefore restrict ourselves to the range of angular exponents

$$\beta = 2,\qquad \alpha \gtrsim 2.\qquad(4.10)$$

We expect that for $\alpha < 2$ our effective field theory description will begin to break down, while as α is increased above 2 it should improve.

4.3 Factorization and Effective Field Theory Analysis

In each region of phase space identified in Sec. 4.2, hierarchies of scales associated with the particular kinematic configuration of the jet appear. These include the soft subjet energy fraction z_{sj} in the soft subjet region of phase space, or the splitting angle θ_{12} of the collinear

subjets. Logarithms of these scales appear at each order in perturbation theory, and need to be resummed to all orders to achieve reliable predictions. To perform this resummation, we will prove factorization theorems in each region of phase space by developing an effective field theory description which captures all the scales relevant to that particular region of phase space. These effective field theories are formulated in the language of SCET [22, 23, 24, 25], but include additional modes which are required to describe the dynamics of the scales associated with the jet's particular substructure. Resummation is then achieved by renormalization group evolution within the effective theory.

In this section we discuss each of the effective theories required for a description of the $e_2^{(\alpha)}, e_2^{(\beta)}, e_3^{(\alpha)}$ phase space. For each region of the phase space, we present an analysis of the modes required in the effective field theory description and present the factorization theorem. We also provide a brief discussion of the physics described by each of the functions appearing in the factorization theorem. Field theoretic operator definitions of the functions, as well as their calculation to one-loop accuracy, are presented in appendices.

4.3.1   QCD Background

Three distinct factorization theorems are required to describe the full phase space for massive QCD jets, corresponding to the soft haze, collinear subjets, and soft subjet configurations. Detailed expositions of the factorization theorems for the collinear subjets and soft subjet configurations have been presented in Refs. [105, 11], but here we review the important features of the factorization theorems to keep the discussion self-contained.

Throughout this section, all jets are defined using the e^+e^- anti-k_T clustering metric [94, 100] with the Winner-Take-All (WTA) recombination scheme [33, 89]. To focus on the aspects of the factorization relevant to the jet substructure, we will present the factorization theorems for the specific case of $e^+e^- \to q\bar{q}$. The factorization theorem for gluon initiated jets is identical to the quark case, and can be performed using the ingredients in the appendices. The extension to the production of additional jets or pp colliders will be discussed in Sec. 4.7.

124

Collinear Subjets

An effective field theory describing the collinear subjets configuration was first presented in Ref. [105] and is referred to as SCET$_+$. We refer the interested reader to Ref. [105] for a more detailed discussion, as well as a formal construction of the effective theory. To our knowledge, our calculation is the first, other than that of Ref. [105], to use this effective theory.

Mode Structure

The modes of SCET$_+$ are global soft modes, two collinear sectors describing the radiation in each of the collinear subjets, and collinear-soft modes from the dipole of the subjet splitting. These are shown schematically in Fig. 4-6. The additional collinear-soft modes, as compared with traditional SCET, are necessary to resum logarithms associated with the subjets' splitting angle. This angle, which is taken to be small, is not resolved by the long wavelength global soft modes.

The parametric scalings of the observables in the collinear subjets region were given in Sec. 4.2.2 and are:

$$e_2^{(\alpha)} \sim \theta_{12}^\alpha \,, \tag{4.11}$$

$$e_2^{(\beta)} \sim \theta_{12}^\beta \,, \tag{4.12}$$

$$e_3^{(\alpha)} \sim \theta_{12}^\alpha (\theta_c^\alpha \theta_{12}^\alpha + z_s + \theta_{12}^{2\alpha} z_{cs}) \,. \tag{4.13}$$

Although the measurement of two 2-point energy correlation functions is required to be able to distinguish the soft and collinear subjets, they are redundant in the collinear subjets region from a power counting perspective, due to the relation $e_2^{(\beta)} \sim \left(e_2^{(\alpha)}\right)^{\beta/\alpha}$. We will therefore always write the scaling of the modes in terms of $e_2^{(\alpha)}$ and $e_3^{(\alpha)}$ to simplify expressions.

From Eq. (4.11), we see that $e_2^{(\alpha)}$ sets the hard splitting scale, while the scalings of all the modes are set by the measurement of $e_3^{(\alpha)}$. In particular, the scaling of the momenta of

the collinear and soft modes are given by

$$p_c \sim E_J \left(\left(\frac{e_3^{(\alpha)}}{\left(e_2^{(\alpha)}\right)^2} \right)^{2/\alpha}, 1, \left(\frac{e_3^{(\alpha)}}{\left(e_2^{(\alpha)}\right)^2} \right)^{1/\alpha} \right)_{n_a \bar{n}_a, n_b \bar{n}_b}, \tag{4.14}$$

$$p_s \sim E_J \frac{e_3^{(\alpha)}}{e_2^{(\alpha)}} (1, 1, 1)_{n\bar{n}}, \tag{4.15}$$

while the scaling of the collinear-soft mode is given by

$$p_{cs} \sim E_J \frac{e_3^{(\alpha)}}{\left(e_2^{(\alpha)}\right)^3} \left(\left(e_2^{(\alpha)}\right)^{2/\alpha}, 1, \left(e_2^{(\alpha)}\right)^{1/\alpha} \right)_{n\bar{n}}. \tag{4.16}$$

Here E_J is the energy of the jet, and the subscripts denote the light-like directions with respect to which the momenta is decomposed. In the expressions above, the momenta are written in the $(+, -, \perp)$ component basis with respect to the appropriate light-like directions. The subjet directions are labelled by n_a and n_b, while the fat jet (containing the two subjets) and the recoiling jet are labelled by n and $\bar{n}$. The relevant modes and a schematic depiction of the hierarchy of their virtualities is shown in Fig. 4-6.

To have a valid soft and collinear expansion, the scalings of the modes in Eqs. (4.14) and (4.16) imply that

$$e_2^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^{\alpha/\beta} \ll 1 \quad \text{and} \quad \frac{e_3^{(\alpha)}}{\left(e_2^{(\alpha)}\right)^3} \sim \frac{e_3^{(\alpha)}}{\left(e_2^{(\beta)}\right)^{3\alpha/\beta}} \ll 1. \tag{4.17}$$

This agrees with the boundaries of the phase space found in Sec. 4.2.2.

Factorization Theorem

In the collinear subjets region of phase space, the values of the 2-point energy correlation functions $e_2^{(\alpha)}$ and $e_2^{(\beta)}$ are set by the hard splitting. To leading power, these observables can be used to provide IRC safe definitions of the subjet energy fractions and the angle between

Collinear Subjets

$$e_3^{(\alpha)} \ll \left(e_2^{(\alpha)}\right)^3$$

$$e_2^{(\beta)} \sim \left(e_2^{(\alpha)}\right)^{\beta/\alpha}$$

J_{n_b}

jet axis

$S^+_{n_a n_b \bar{n}_t}$

R

J_{n_a}

$S_{\bar{n}_t \bar{n}_t}$

Match To SCET

Factorize Jet Function

H H $\mu \sim E_J$

$J_n, J_{\bar{n}}$ $P_{n_t \to n_a, n_b}$ $\mu \sim E_J e_2^{(2)}$

$J_{n_a}, J_{n_b}, J_{\bar{n}_t}$ $\mu \sim E_J \left(\dfrac{e_3^{(2)}}{\left(e_2^{(2)}\right)^2}\right)^{1/2}$

$S_{n\bar{n}}$ $S^+_{n_a n_b \bar{n}}$ $\mu \sim E_J \dfrac{e_3^{(2)}}{\left(e_2^{(2)}\right)^3}\left(e_2^{(2)}\right)^{(1/2)}$

$S_{n_t \bar{n}_t}$ $\mu \sim E_J \dfrac{e_3^{(2)}}{e_2^{(2)}}$

Increasing Virtuality

(a)   (b)

Figure 4-6: A schematic depiction of the collinear subjets configuration with dominant QCD radiation and the functions describing its dynamics in the effective field theory is shown in a). The matching procedure and relevant scales are shown in b), where we have restricted to the case $\alpha = \beta = 2$ for simplicity.

the subjets. We therefore write the factorization theorem in terms of $e_2^{(\alpha)}$, $e_3^{(\alpha)}$, and the energy fraction of one of the subjets, which we denote by z. We further assume that an IRC safe observable, B, is measured in the out-of-jet region. Dependence on B enters only into the out-of-jet jet function, and the out-of-jet contribution to the soft function.

The factorization theorem formulated in SCET$_+$ for the collinear subjets region of phase space is given by

$$\frac{d^3\sigma}{dz\, de_2^{(\alpha)}\, de_3^{(\alpha)}} = \sum_{f,f_a,f_b} H^f_{n\bar{n}} J_{\bar{n}}(B) P^{f \to f_a f_b}_{n_t \to n_a, n_b}\left(z; e_2^{(\alpha)}\right) \int de_3^c de_3^{\bar{c}} de_3^s de_3^{cs} \tag{4.18}$$

$$\times \delta\left(e_3^{(\alpha)} - e_3^c - e_3^{\bar{c}} - e_3^s - e_3^{cs}\right) J^{f_a}_{n_a}\left(z; e_3^c\right) J^{f_b}_{n_b}\left(1 - z; e_3^{\bar{c}}\right) S_{n\bar{n}}\left(e_3^s, B; R\right) S^+_{n_a n_b \bar{n}}\left(e_3^{cs}\right),$$

where we have suppressed the convolution over the out-of-jet measurement, B, for simplicity. Here the n_a, n_b denote the collinear directions of the subjets, and we assume that $z \sim 1 - z \sim \frac{1}{2}$. The sum runs over all possible quark flavors that could be produced in an e^+e^- collision. A brief description of the functions entering the factorization theorem of Eq. (4.18) is as

127

follows:

- $H_{n\bar{n}}^{f}$ is the hard function describing the underlying short distance process. In this case we consider $e^{+}e^{-} \to q\bar{q}$.

- $P_{n\to n_a,n_b}^{f\to f_a f_b}\left(z; e_2^{(\alpha)}\right)$ is the hard function arising from the matching for the hard splitting into subjets. In this case the partonic channel $f \to f_a f_b$ is restricted to $q \to qg$.

- $J_{n_a}^{f_a}\left(z; e_3^c\right)$, $J_{n_b}^{f_b}\left(1 - z; e_3^{\bar{c}}\right)$ are jet functions describing the collinear dynamics of the subjets along the directions n_a, n_b.

- $S_{n\bar{n}}(e_3^s, B; R)$ is the global soft function. The global soft modes do not resolve the subjet splitting, and are sensitive only to two eikonal lines in the n and $\bar{n}$ directions. The soft function depends explicitly on the jet radius, R.

- $S_{n_a n_b \bar{n}}^{+}\left(e_3^{cs}\right)$ is the collinear-soft function. The collinear-soft modes resolve the subjet splitting, and hence the function depends on three eikonal lines, namely $n_a, n_b, \bar{n}$. Although these modes are soft, they are also boosted, and therefore do not resolve the jet boundary, so that the collinear soft function is independent of the jet radius, R.

This factorization theorem is shown schematically in Fig. 4-6, which highlights the radiation described by each of the functions in Eq. (4.18), as well as their virtuality scales. The two stage matching procedure onto the SCET$_{+}$ effective theory, which proceeds through a refactorization of the jet function, is also shown. The fact that the refactorization occurs in the jet function is important in that it implies that it is independent of the global color structure of the event, making it trivial to extend the factorization theorem to events with additional jets. This matching procedure is discussed in detail in Ref. [105].

Operator definitions, and one-loop calculations for the operators appearing in the factorization theorem of Eq. (4.18) are given in App. A.2.
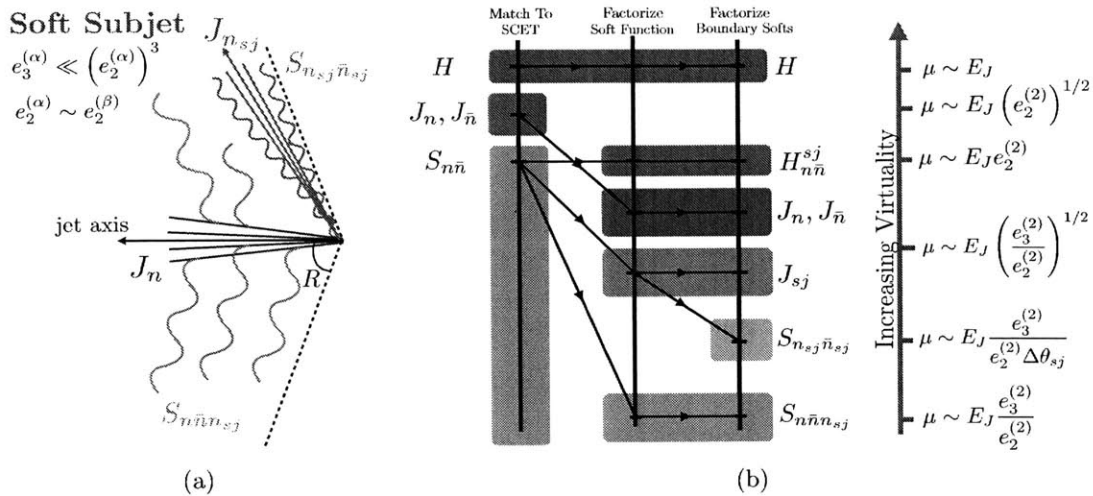
Figure 4-7: A schematic depiction of the soft subjet configuration with dominant QCD radiation and the functions describing its dynamics in the effective field theory is shown in a). The matching procedure and relevant scales are shown in b), where we have restricted to the case $\alpha = \beta = 2$ for simplicity.

Soft Subjet

A factorization theorem describing the soft subjet region of phase space was recently presented in Ref. [11]. In this section we review the basic features of this factorization theorem, but we refer the reader to Ref. [11] for a more detailed discussion.

Unlike for the case of collinear subjets, in the soft subjet configuration, the wide angle soft subjet probes the boundary of the jet. This introduces sensitivity to the details of the jet algorithm used to define the jet, as well as to the measurement made in the region outside the jet. The factorization theorem of Ref. [11] is valid under the assumption that an additive IRC safe observable, B, is measured in the out-of-jet region, and that the soft scale associated with this observable, Λ, satisfies $\Lambda / E_J \ll e_2^{(\alpha)}$. We will therefore assume that this condition is satisfied throughout this section. However, we will see that the numerical results are fairly insensitive to the details of the choice of scale Λ. Ref. [11] also used a broadening axis [33] cone algorithm to define jets, whereas here we use the anti-k_T algorithm, as relevant for phenomenological applications. We will argue that the structure of the factorization theorem

129

is in fact identical in the two cases, to leading power.

Mode Structure

In the soft subjet region of phase space there are two subjets with an energy hierarchy. We denote the energy of the soft subjet by z_{sj} and the angle from the n axis by θ_{sj}. We also use the notation $\Delta\theta_{sj} = R - \theta_{sj}$ to denote the angle from the soft subjet axis to the jet boundary. The modes of the soft subjet are collinear-soft modes, being both soft and collimated, and we will therefore denote the characteristic angle between them as θ_{cs}. Straightforward power counting can be applied to determine the scaling of the modes for both the energetic jet and the soft subjet. Their contributions to the observable are given by

$$e_2^{(\alpha)} \sim z_{sj} \,, \tag{4.19}$$

$$e_2^{(\beta)} \sim z_{sj} \,, \tag{4.20}$$

$$e_3^{(\alpha)} \sim z_{sj}(\theta_c^\alpha + z_{sj}\theta_{cs}^\alpha + z_s) \,. \tag{4.21}$$

In the soft subjet region of phase space, we have the relation $e_2^{(\alpha)} \sim e_2^{(\beta)}$, and therefore these two observables are redundant from a power counting perspective. We will therefore write the power counting of the modes in terms of $e_2^{(\alpha)}$ and $e_3^{(\alpha)}$.

From the contributions to the observables above, we find that the momentum of the collinear and global soft radiation scales like

$$p_c \sim E_J \left(\left(\frac{e_3^{(\alpha)}}{e_2^{(\alpha)}} \right)^{2/\alpha}, 1, \left(\frac{e_3^{(\alpha)}}{e_2^{(\alpha)}} \right)^{1/\alpha} \right)_{n\bar{n}} \,, \tag{4.22}$$

$$p_s \sim E_J \frac{e_3^{(\alpha)}}{e_2^{(\alpha)}} (1,1,1)_{n\bar{n}} \,,$$

where E_J is the energy of the jet and n and $\bar{n}$ are the light-like directions of the jet of interest

and the other jet in the event, respectively. The soft subjet mode's momentum scales like

$$p_{sj} \sim E_J \, e_2^{(\alpha)} \left(\left(\frac{e_3^{(\alpha)}}{\left(e_2^{(\alpha)}\right)^2} \right)^{2/\alpha} , 1, \left(\frac{e_3^{(\alpha)}}{\left(e_2^{(\alpha)}\right)^2} \right)^{1/\alpha} \right)_{n_{sj} \bar{n}_{sj}} , \qquad (4.23)$$

in the light-cone coordinates defined by the direction of the soft subjet, n_{sj}. These are the complete set of modes defined by the scales set by the measurements of $e_2^{(\alpha)}, e_2^{(\beta)}$, and $e_3^{(\alpha)}$ alone.

Unlike in the collinear subjet region of phase space there are no collinear-soft modes required in the effective field theory description, since the soft subjet is at a wide angle from the jet axis. However, in this region there is an additional mode, termed a *boundary soft mode* in Ref. [11], whose appearance is forced by the jet boundary and the energy veto in the region of phase space outside the jet. These modes do not contribute to the e_2 observables, but are effectively a collinear-soft mode whose angle with respect to the soft subjet axis is set by the angle to the boundary. The boundary soft mode's momentum components scale like

$$p_{bs} \sim E_J \frac{e_3^{(\alpha)}}{e_2^{(\alpha)} \left(\Delta\theta_{sj}\right)^\alpha} \left(\left(\Delta\theta_{sj}\right)^2, 1, \Delta\theta_{sj} \right)_{n_{sj}\bar{n}_{sj}} , \qquad (4.24)$$

written in the light-cone coordinates defined by the soft subjet axis. The boundary soft modes are required to have a single scale in the soft subjet function. For consistency of the factorization, we must enforce that the soft subjet modes cannot resolve the jet boundary and that the boundary soft modes are localized near the jet boundary. That is, the angular size of the soft subjet modes, θ_{cs}, must be parametrically smaller than that of the boundary soft modes, namely $\Delta\theta_{sj}$. We therefore find the condition

$$\left(\Delta\theta_{sj}\right)^\alpha \gg \left(\theta_{cs}\right)^\alpha \sim \frac{e_3^{(\alpha)}}{\left(e_2^{(\alpha)}\right)^2} , \qquad \text{and} \qquad \Delta\theta_{sj} \ll 1 . \qquad (4.25)$$

131

Therefore, the factorization theorem applies in a region of the phase space where the soft subjet is becoming pinched against the boundary of the jet, but lies far enough away that the collinear modes of the soft subjet do not touch the boundary. A schematic depiction of this region of phase space, along with a summary of all the relevant modes which appear in the factorization theorem is shown in Fig. 4-7.

In the soft subjet region of phase space, the choice of jet algorithm plays a crucial role, since the soft subjet probes the boundary of the jet. In Ref. [11] the factorization theorem in the soft subjet region of phase space was presented using a broadening axis cone algorithm with radius R. We now show that up to power corrections, the factorization theorem in the soft subjet region of phase space is identical with either the anti-k_T or broadening axis cone algorithm. In particular, with the anti-k_T algorithm, the jet boundary is not deformed by the soft subjet, and can be treated as a fixed cone of radius R. This is not true for other jet algorithms, such as such as k_T [158, 159] or Cambridge-Aachen [160, 161, 162], where the boundary is deformed by the clustering of soft emissions, a point which has been emphasized elsewhere (see, e.g., Refs. [163, 164, 165, 166]).

The validity of the factorization theorem requires the following two conditions, which will put constraints on the power counting in the soft subjet region of phase space. First, the soft subjet must be clustered with the jet axis, rather than with the out-of-jet radiation. This is guaranteed as long as the soft subjet axis satisfies $\theta_{sj} < R$. Second, the radiation clustered with the soft subjet from the out-of-jet region should not distort the boundary of the jet. More precisely, the distortion of the boundary must not modify the value of $e_3^{(\alpha)}$ at leading power (note that the power counting guarantees that it does not modify $e_2^{(\alpha)}$). The contribution to $e_3^{(\alpha)}$ from a soft out-of-jet emission is given by

$$e_2^{(\alpha)} \frac{\Lambda}{E_J} \ll e_3^{(\alpha)} \implies \frac{\Lambda}{E_J} \ll \frac{e_3^{(\alpha)}}{e_2^{(\alpha)}} \sim \left(e_2^{(\alpha)} \right)^2. \tag{4.26}$$

Since the out-of-jet scale is in principle a free parameter, we can formally enforce this condition in our calculations. Corrections due to a deformation of the jet boundary would enter as power corrections in this region of phase space. The jet boundary therefore acts as a

hard boundary of radius R, and the factorization theorem is identical to that presented in Ref. [11].

Factorization Theorem

With an understanding of the precise restrictions on the power counting required for the validity of the soft subjet factorization theorem, we now discuss its structure. Since we have argued that the relevant factorization theorem is identical to that presented in Ref. [11], we will only state the result. The factorization theorem in the soft subjet region with the out-of-jet scale satisfying $\Lambda \ll e_2^{(\alpha)} E_J$, and with jets defined by the anti-k_T jet algorithm, is given by

$$\frac{d\sigma(B;R)}{de_2^{(\alpha)} de_2^{(\beta)} de_3^{(\alpha)}} = \tag{4.27}$$

$$\int dB_S dB_{J_{\bar{n}}} \int de_3^{J_n} de_3^{J_{sj}} de_3^S de_3^{S_{sj}} \delta(B - B_{J_{\bar{n}}} - B_S) \delta(e_3^{(\alpha)} - e_3^{J_n} - e_3^{J_{sj}} - e_3^S - e_3^{S_{sj}})$$

$$\times H_{n\bar{n}}(Q^2) H_{n\bar{n}}^{sj}\left(e_2^{(\alpha)}, e_2^{(\beta)}\right) J_n\left(e_3^{J_n}\right) J_{\bar{n}}(B_{J_{\bar{n}}}) S_{n\bar{n}n_{sj}}\left(e_3^S; B_S; R\right) J_{n_{sj}}\left(e_3^{J_{sj}}\right) S_{n_{sj}\bar{n}_{sj}}\left(e_3^{S_{sj}}; R\right).$$

In this expression we have explicitly indicated the dependence on the jet boundaries with the jet radius R. A brief description of the functions appearing in Eq. (4.27) is as follows:

- $H_{n\bar{n}}(Q^2)$ is the hard function describing the underlying short distance process. In this case we consider $e^+ e^- \to q\bar{q}$.

- $H_{n\bar{n}}^{sj}\left(e_2^{(\alpha)}, e_2^{(\beta)}\right)$ is the hard function describing the production of the soft subjet coherently from the initial $q\bar{q}$ dipole, and describes dynamics at the scale set by $e_2^{(\alpha)}, e_2^{(\beta)}$.

- $J_n\left(e_3^{(\alpha)}\right)$ is a jet function at the scale $e_3^{(\alpha)}$ describing the hard collinear modes of the identified jet along the n direction.

- $J_{\bar{n}}(B)$ is a jet function describing the collinear modes of the out-of-jet region of the event.

133

- $S_{n\bar{n}n_{sj}}\left(e_3^{(\alpha)}; B; R\right)$ is the global soft function involving three Wilson line directions, $n, \bar{n}, n_{sj}$. The global soft function depends explicitly on both the out-of-jet measurement and the jet radius.

- $J_{n_{sj}}\left(e_3^{(\alpha)}\right)$ is a jet function describing the dynamics of the soft subjet modes, which carry the bulk of the energy in the soft subjet.

- $S_{n_{sj}\bar{n}_{sj}}(e_3^{(\alpha)}; R)$ is a soft function describing the dynamics of the boundary soft modes. It depends only on two Wilson line directions $n_{sj}, \bar{n}_{sj}$.

These functions, and a schematic depiction of the radiation which they define, are indicated in Fig. 4-7, along with a schematic depiction of the multistage matching procedure from QCD onto the effective theory, as described in detail in Ref. [11]. Although we will not discuss any details of the matching procedure, it is important to note that it occurs through a refactorization of the soft function, and hence the soft subjet factorization theorem is sensitive to the global color structure of the event, since the soft subjet is emitted coherently from all eikonal lines. This should be contrasted with the case of the collinear subjets factorization theorem, where the matching occurs through a refactorization of the jet function.

In the soft subjet region of phase space, we can relate the variables $e_2^{(\alpha)}, e_2^{(\beta)}$ to the physically more transparent z_{sj}, θ_{sj} variables with a simple Jacobian factor, giving the factorization theorem

$$\frac{d\sigma(B; R)}{dz_{sj}\, d\theta_{sj}\, de_3^{(\alpha)}} = \tag{4.28}$$

$$\int dB_S dB_{J_{\bar{n}}} \int de_3^{J_n} de_3^{J_{sj}} de_3^S de_3^{S_{sj}} \delta(B - B_{J_{\bar{n}}} - B_S)\delta(e_3^{(\alpha)} - e_3^{J_n} - e_3^{J_{sj}} - e_3^S - e_3^{S_{sj}})$$

$$\times H_{n\bar{n}}(Q^2) H_{n\bar{n}}^{sj}\left(z_{sj}, \theta_{sj}\right) J_n\left(e_3^{J_n}\right) J_{\bar{n}}(B_{J_{\bar{n}}}) S_{n\bar{n}n_{sj}}\left(e_3^S; B_S; R\right) J_{n_{sj}}\left(e_3^{J_{sj}}\right) S_{n_{sj}\bar{n}_{sj}}(e_3^{S_{sj}}; R).$$

Operator definitions, and one-loop calculations for the operators appearing in the factorization theorem of Eqs. (4.27) and (4.28) are given in App. A.3.

Soft Haze

$$e_3^{(\alpha)} \sim \left(e_2^{(\alpha)}\right)^2$$
$$e_2^{(\alpha)} \sim e_2^{(\beta)}$$

jet axis

J_n

R

$S_{n\bar{n}}$

(a)

Match To SCET

H

$J_n, J_{\bar{n}}$

$S_{n\bar{n}}$

Increasing Virtuality

$\mu \sim E_J$

$\mu \sim E_J \left(e_2^{(2)}\right)^{1/2}$

$\mu \sim E_J e_2^{(2)}$

(b)
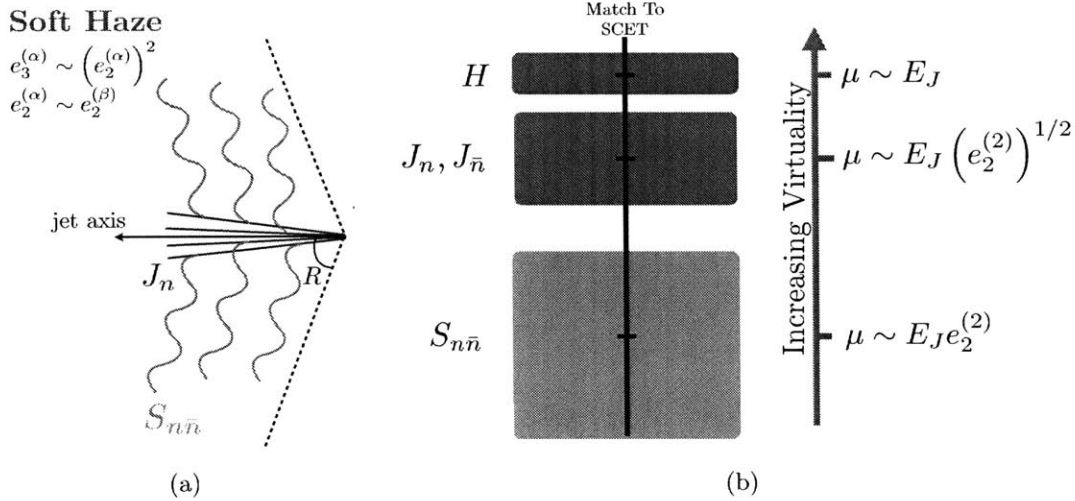
Figure 4-8: A schematic depiction of the soft haze configuration where no subjets are resolved, with dominant QCD radiation and the functions describing its dynamics in the effective field theory is shown in a). The relevant scales in the effective field theory are shown in b), where we have restricted to the case $\alpha = \beta = 2$ for simplicity.

Soft Haze

The soft haze region defines the upper boundary of the $e_2^{(\beta)}, e_3^{(\alpha)}$ phase space. In this region of phase space jets consist of a single hard core, with no resolved subjets. A factorization theorem describing this region of phase space has not been presented elsewhere, but can be straightforwardly formulated in standard SCET involving only n and $\bar{n}$ collinear sectors.

As discussed in Sec. 4.2.2, the power counting in the soft haze region depends sensitively on the relative values of α and β, and therefore so does the structure of the factorization theorem. Since, from Eq. (4.10), we restrict ourself to $\alpha \geq \beta$, we will for simplicity only discuss the factorization theorems valid in this case. Factorization theorems for other values of α and β can be determined by performing a similar analysis.

135

Mode Structure

In the soft haze region the observables have the power counting

$$e_2^{(\alpha)} \sim z_s + \theta_c^\alpha \,, \tag{4.29}$$

$$e_2^{(\beta)} \sim z_s + \theta_c^\beta \,, \tag{4.30}$$

$$e_3^{(\alpha)} \sim z_s^2 + \theta_c^\alpha z_s + \theta_c^{3\alpha} \,, \tag{4.31}$$

where we have not yet dropped power suppressed terms. We are interested in the factorization theorem on the upper boundary, with the scaling $e_3^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^2$.[4] We now assume $\alpha > \beta$. In this case, dropping power suppressed terms, the appropriate power counting is

$$e_2^{(\alpha)} \sim z_s \,, \tag{4.32}$$

$$e_2^{(\beta)} \sim z_s + \theta_c^\beta \,, \tag{4.33}$$

$$e_3^{(\alpha)} \sim z_s^2 \,. \tag{4.34}$$

It is also interesting to consider the case $\alpha = \beta$ because in the soft haze region it is not necessary to measure two different 2-point energy correlation functions, unlike in the two-prong region of phase space. In the case that $\alpha = \beta$, we have instead,

$$e_2^{(\alpha)} \sim z_s + \theta_c^\alpha \,, \tag{4.35}$$

$$e_3^{(\alpha)} \sim z_s^2 + \theta_c^\alpha z_s \,, \tag{4.36}$$

---

[4]There is another parametric choice for the relative scaling of the 2-point energy correlation functions [10], though it does not extend to the upper boundary of the phase space. If $(e_2^{(\alpha)})^\beta \sim (e_2^{(\beta)})^\alpha$, then the power counting is

$$e_2^{(\alpha)} \sim z_s + \theta_c^\alpha \,,$$
$$e_2^{(\beta)} \sim \theta_c^\beta \,,$$
$$e_3^{(\alpha)} \sim z_s^2 + \theta_c^\alpha z_s \,,$$

with both 2-point correlation functions dominated by collinear physics. For $\alpha > \beta$, this region has the scaling $e_3^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^{2\alpha/\beta}$ which does not extend to the upper boundary.

136

where the second term in the expression for $e_3^{(\alpha)}$ is no longer power suppressed. This will modify the factorization theorem between the two cases.

In both cases, the scaling of the modes is then given by

$$p_c \sim E_J \left(\left(e_2^{(\beta)} \right)^{2/\beta}, 1, \left(e_2^{(\beta)} \right)^{1/\beta} \right)_{n\bar{n}}, \tag{4.37}$$

$$p_s \sim e_2^{(\beta)} E_J \left(1, 1, 1 \right)_{n\bar{n}}, \tag{4.38}$$

with $\beta = \alpha$ in the second case. Here E_J is the energy of the jet and the subscripts denote the light-like directions with respect to which the momenta is decomposed. This scaling should be recognized as the usual power counting of the collinear and soft modes for the angularities with angular exponent β [167, 10].

Factorization Theorem

The factorization theorem in the soft haze region of phase space can now be straightforwardly read off from the power counting expressions of the previous sections. We state it both for the case $\alpha = \beta$ and $\alpha > \beta$. For $\alpha > \beta$, we have

$$\frac{d\sigma}{de_2^{(\alpha)} de_2^{(\beta)} de_3^{(\alpha)}} = H_{n\bar{n}}(Q^2) J_{\bar{n}}(B) \int de_2^c de_2^s \delta \left(e_2^{(\beta)} - e_2^c - e_2^s \right) J_n \left(e_2^c \right) S_{n\bar{n}} \left(e_2^s, e_2^{(\alpha)}, e_3^{(\alpha)}, R, B \right),$$

$$\tag{4.39}$$

where we have suppressed the convolution over the out-of-jet measurement B, to focus on the structure of the in-jet measurements. For $\alpha = \beta$, the factorization theorem takes an

137

interesting form[5]

$$\frac{d\sigma}{de_2^{(\alpha)} de_3^{(\alpha)}} = H_{n\bar{n}}(Q^2) J_{\bar{n}}(B) \int de_2^c de_2^s de_3^s \delta \left(e_2^{(\alpha)} - e_2^c - e_2^s \right) \delta \left(e_3^{(\alpha)} - e_2^c e_2^s - e_3^s \right) \quad (4.40)$$

$$\times J_n \left(e_2^c \right) S_{n\bar{n}} \left(e_2^s, e_3^s, R, B \right),$$

where again the convolution over B has been suppressed. A brief description of the functions appearing in the factorization theorems is as follows:

- $H_{n\bar{n}} (Q^2)$ is the hard function describing the underlying short distance process. In this case we consider $e^+ e^- \to q\bar{q}$.

- $J_{\bar{n}}(B)$ is the jet function describing the collinear modes for the recoiling jet.

- $J_n \left(e_2^c \right)$ is the jet function describing the collinear modes for the jet in the n direction.

- $S_{n\bar{n}} \left(e_2^s, e_3^s, R, B \right)$ and $S_{n\bar{n}} \left(e_2^s, e_2^{(\alpha)}, e_3^{(\alpha)}, R, B \right)$ are soft functions describing the global soft radiation from the $n\bar{n}$ dipole. These also carry the jet algorithm constraints denoted by R, and any out-of-jet measurements B.

These functions, and a schematic depiction of the radiation which they define are indicated in Fig. 4-8. In App. A.6, we give operator definitions of these functions and the leading-power expression for the $e_3^{(\alpha)}$ measurement operator in the soft function.

There are several interesting features about the factorization theorems of Eqs. (4.39) and (4.40). First, the soft functions are multi-differential, in that they require the simultaneous measurement of multiple quantities. Such multi-differential jet and soft functions have been discussed in detail in Ref. [10, 34]. One other interesting feature of the factorization theorem of Eq. (4.40), for the case of equal angular exponents, is the appearance of the product structure in the δ-function defining the value of $e_3^{(\alpha)}$. This product structure is consistent

---

[5]When calculating the tail of the D_2 distribution, one might be tempted to marginalize over $e_2^{(\beta)}$ in Eq. (4.39). This naïve marginalization does not yield the correct result. Rather, if one started the derivation of the factorization theorem with only the measurements of $e_2^{(\alpha)}$ and $e_3^{(\alpha)}$ imposed, so that all possible $e_2^{(\beta)}$ configurations are integrated over, then Eq. (4.40) would be obtained. Thus Eq. (4.40) is the correct marginalization over $e_2^{(\beta)}$ in Eq. (4.39).

with the power counting of Eq. (4.35) which describes the properties of the 3-point energy correlation function in the soft and collinear limits. It is important to note that this product form does not violate soft-collinear factorization, since only the knowledge of the total $e_2^{(\alpha)}$ of the soft or collinear sector is required.

The soft contribution to the 3-point energy correlation is first non-vanishing with two real emissions. Therefore at one-loop, the factorization theorem of Eq. (4.39) reduces exactly to the factorization theorem for the multi-differential angularities studied in Refs. [10, 34], whereas the factorization theorem of Eq. (4.40) reduces to the factorization theorem for a single angularity. In this chapter, we will not perform the two-loop calculation necessary to obtain a non-trivial contribution to the three point energy correlation function. Instead, we will obtain an approximation to the cross section in this region by taking a limit of our factorization theorems in the two-prong region of phase space. This is possible, because as we will show in Sec. 4.4.3 by studying the fixed order distributions for the observable D_2, there is no fixed order singularity in the soft haze region of phase space in the presence of a mass cut. This implies that the resummation is not needed to regulate a fixed order singularity. This will be discussed in Sec. 4.4.4. The field theoretic definitions of the functions appearing in the factorization theorem of Eq. (4.39) as well as power expansions of the measurement operators are collected in App. A.6. However, because of the fact that we do not explicitly use the results of the soft haze factorization theorem in our calculation, we simply refer the reader to Refs. [10, 34] for the calculations of the one-loop functions relevant to the factorization theorems of Eqs. (4.39) and (4.40), and leave for future work the full two-loop calculation.

Refactorization of the Global Soft Function

In each of the factorization theorems required for the description of QCD background jets, namely the collinear subjets, soft subjet, and soft haze factorization theorems, there is a global soft function, which is sensitive to both the in-jet measurement of the energy correlation functions, as well as the out-of-jet measurement B. To ensure that all large logarithms

139

are resummed by the renormalization group evolution, we must perform a refactorization of the soft function [168, 167, 169, 82, 84]. This ensures that the only logarithms which appear in a given soft function that are sensitive to both in-jet and out-of-jet scales are true non-global logarithms (NGLs) [38], which first appear at two-loop order in the calculation of a particular soft function.[6] Here we focus on the refactorization of the soft subjet and collinear subjets factorization theorems of Secs. 4.3.1 and 4.3.1, which will be used in our numerical calculation. For both of these factorization theorems, we can write the soft function to all orders in α_s as

$$
S\left(e_3^{(\alpha)}, B; R, \mu\right) = S^{(\text{out})}\left(B; R, \mu\right) S^{(\text{in})}\left(e_3^{(\alpha)}; R, \mu\right) S_{\text{NGL}}\left(e_3^{(\alpha)}, B; R\right) , \qquad (4.41)
$$

where we have explicitly indicated the renormalization scale μ dependence [170]. The non-global part of the soft function $S_{\text{NGL}}\left(e_3^{(\alpha)}, B; R\right)$ is first non-trivial at two-loop order, beyond the accuracy to which we explicitly calculated the soft functions in this chapter. Furthermore, the anomalous dimension of the soft function factorizes to all orders in perturbation theory as

$$
\gamma_S\left(e_3^{(\alpha)}, B; R; \mu\right) = \gamma_S^{(\text{out})}\left(B; R; \mu\right) + \gamma_S^{(\text{in})}\left(e_3^{(\alpha)}; R; \mu\right) , \qquad (4.42)
$$

and therefore the renormalization group kernels factorize as well. Briefly, this occurs because renormalization group consistency relates the soft anomalous dimension to the sum of all the other anomalous dimensions, each of which can be associated with the in-jet or out-of-jet contributions.[7]

While similar refactorizations of the global soft function have been discussed previously,

---

[6] It is important to emphasize that throughout this section we refer to the NGLs which appear in the soft function of a given factorization theorem, and the order in α_s at which they will appear in this particular soft function. Because we combine distinct factorization theorems, some of which include hard splitting functions, or eikonal emission functions, this order is in general distinct from the order at which they will appear in the total cross section, which can be different for each factorization theorem. This combination of the factorization theorems is completely independent from the refactorization of the soft function in a particular factorization theorem.

[7] As discussed in Ref. |84| there is some ambiguity in how the hard function, for example, is associated with the in-jet or out-of-jet anomalous dimensions, but this does not affect the above argument.

and used in numerical calculations (see especially Ref. [84] for a detailed discussion), we will discuss it here for completeness. The refactorization of the global soft function plays a role in our numerical results and is particularly important in appropriately separating scales in the global soft function of the soft subjet factorization theorem of Sec. 4.3.1. In Ref. [11] the structure of the one-loop calculation of the soft subjet factorization theorem was discussed in detail, with a particular focus on the dependence on the angle $\Delta\theta_{sj}$ between the soft subjet and the boundary. There it was found that the while the out-of-jet soft function contained dependence on the angle between the soft subjet and the boundary, $\Delta\theta_{sj}$, this dependence vanishes in the in-jet contribution to the soft function due to a zero bin subtraction. Renormalization group consistency is achieved since the $\Delta\theta_{sj}$ dependence associated with the in-jet region is carried by the boundary soft function. Therefore, the refactorization of the global soft function for the soft subjet factorization theorem allows the soft function to be separated into a piece with $\Delta\theta_{sj}$ dependence, and a piece with no $\Delta\theta_{sj}$ dependence, and is crucial for resumming all large logarithms associated with this scale. The one-loop anomalous dimensions, split into out-of-jet and in-jet contributions, as well as canonical scales for both the in-jet and out-of-jet soft functions are given in App. A.2, App. A.3, and App. A.4. Further details of this refactorization, and in particular a discussion on the dependence on $\Delta\theta_{sj}$ is also given.

For completeness, we also give the final refactorized expressions for the factorization theorems for the collinear subjets and soft subjet factorization theorems that will be used when presenting numerical results. For the collinear subjets factorization theorem, we have

$$
\frac{d^3\sigma}{dz\,de_2^{(\alpha)}de_3^{(\alpha)}} = \sum_{f,f_a,f_b} H_{n\bar{n}}^f(Q^2) P_{n_t \to n_a,n_b}^{f \to f_a f_b}\left(z; e_2^{(\alpha)}\right) \int dB_S dB_{J_{\bar{n}}} \int de_3^c de_3^{\bar{c}} de_3^s de_3^{cs} \tag{4.43}
$$

$$
\times\, \delta(B - B_{J_{\bar{n}}} - B_S)\delta\left(e_3^{(\alpha)} - e_3^c - e_3^{\bar{c}} - e_3^s - e_3^{cs}\right)
$$

$$
\times\, J_{\bar{n}}(B_{J_{\bar{n}}}) J_{n_a}^{f_a}\left(z; e_3^c\right) J_{n_b}^{f_b}\left(1 - z; e_3^{\bar{c}}\right) S_{n\bar{n}}^{(\text{out})}\left(B_S; R\right) S_{n\bar{n}}^{(\text{in})}\left(e_3^s; R\right) S_{n_a n_b \bar{n}}^+\left(e_3^{cs}\right),
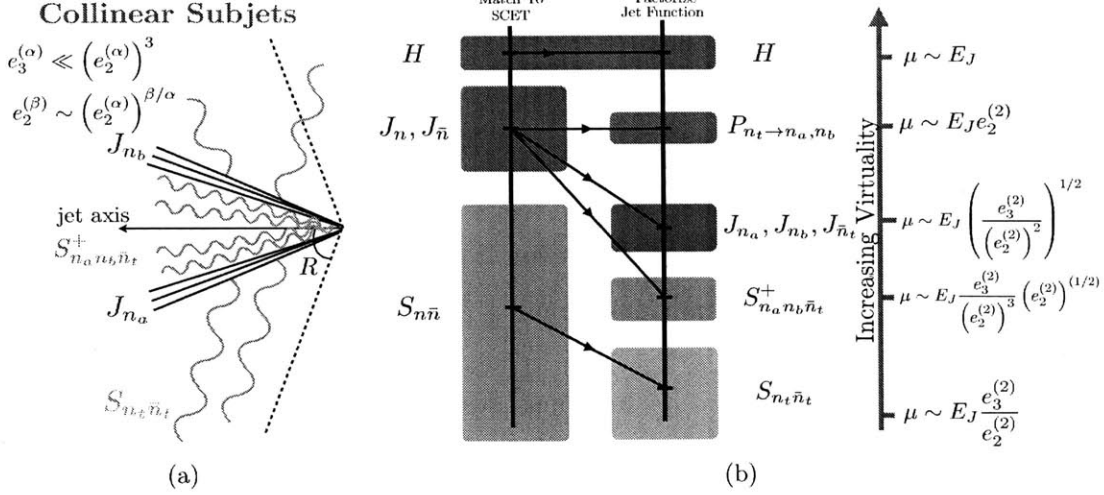$$
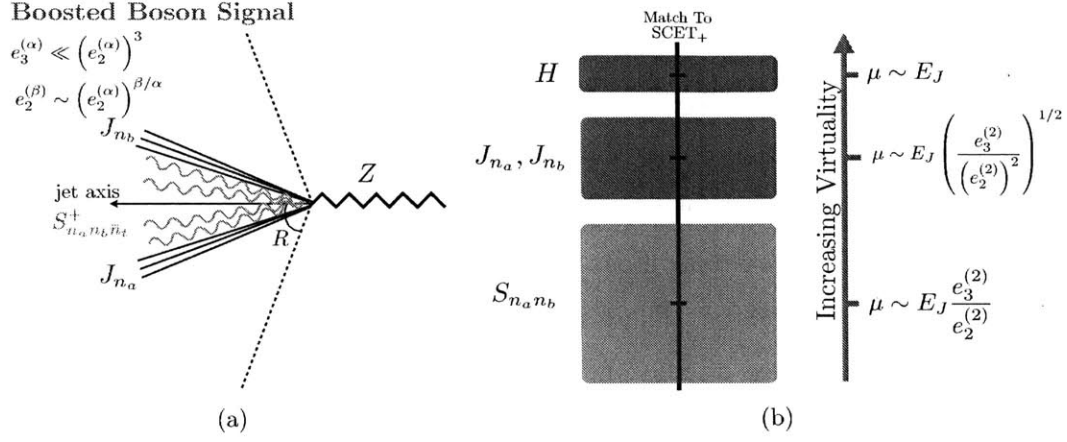
Figure 4-9: A schematic depiction of the boosted Z boson configuration with dominant QCD radiation and the functions describing its dynamics in the effective field theory is shown in a). The relevant scales, ordered in virtuality, are summarized in b), where we have restricted to the case $\alpha = \beta = 2$ for simplicity.

while for the soft subjet factorization theorem, we have

$$\frac{d\sigma(B;R)}{dz_{sj}\,d\theta_{sj}\,de_3^{(\alpha)}} = H_{n\bar{n}}(Q^2)H_{n\bar{n}}^{sj}(z_{sj},\theta_{sj}) \int dB_S dB_{J_{\bar{n}}} \int de_3^{J_n} de_3^{J_{sj}} de_3^S de_3^{S_{sj}} \tag{4.44}$$

$$\times \delta(B - B_{J_{\bar{n}}} - B_S)\delta(e_3^{(\alpha)} - e_3^{J_n} - e_3^{J_{sj}} - e_3^S - e_3^{S_{sj}})$$

$$\times J_n\left(e_3^{J_n}\right) J_{\bar{n}}(B_{J_{\bar{n}}}) S_{n\bar{n}n_{sj}}^{\text{(out)}}\left(B_S;R\right) S_{n\bar{n}n_{sj}}^{\text{(in)}}\left(e_3^S;R\right) J_{n_{sj}}\left(e_3^{J_{sj}}\right) S_{n_{sj}\bar{n}_{sj}}(e_3^{S_{sj}};R).$$

In this form, each function in Eqs. (4.43) and (4.44) contains logarithms of a single scale, which can be resummed through renormalization group evolution.

4.3.2 Boosted Boson Signal

In this section we discuss the effective field theory and factorization theorem relevant for the hadronically-decaying boosted boson signal. For concreteness, we will consider the case of a boosted Z boson decaying to a massless $q\bar{q}$ pair; however, the extension to other color-neutral boosted particles is trivial. We will work in the narrow width approximation, setting the width of the Z boson $\Gamma_Z = 0$. Corrections to this approximation are trivial to implement,

as they do not modify the structure of the factorization, and are expected to have a minimal effect.

A factorization theorem for the N-subjettiness observable $\tau_{1,2}^{(\beta)}$ [104, 32, 31] measured on boosted Z jets was presented in Ref. [85]. This factorization theorem was obtained by boosting an appropriately chosen e^+e^- event shape. A factorization theorem can also be formulated using the SCET$_+$ effective theory,[8] where the collinear-soft mode, which was described in Sec. 4.3.1, corresponds to the boosted soft mode of the e^+e^- event shape. We will take this second approach, as it is in line with the general spirit of this chapter, of developing effective field theory descriptions of jet substructure configurations. However, the approach of relating to boosted e^+e^- event shape variables is useful for relating results to higher order calculations known in the literature. Despite the fact that the factorization for the energy correlation functions in the signal region follows straightforwardly from that of Ref. [85], or from the SCET$_+$ factorization theorem of Sec. 4.3.1, we will discuss it here for completeness.

We assume the process $e^+e^- \to ZZ \to q\bar{q}l\bar{l}$, where l is a lepton to avoid having to describe additional jets, although the extension to two hadronically-decaying Z bosons is trivial. The factorization theorem is then similar to that presented in Sec. 4.3.1, however, there are no global soft modes since the Z is a color singlet. The scaling of the collinear and collinear-soft modes are identical to those given in Sec. 4.3.1, so we do not repeat them here. The factorization theorem is given by

$$\frac{d\sigma}{dz\,de_2^{(\alpha)}de_3^{(\alpha)}} = H(Q^2)P_{n\to n_a,n_b}^{Z\to q\bar{q}}\left(z; e_2^{(\alpha)}\right) \int de_3^c de_3^{\bar{c}} de_3^s de_3^{cs} \tag{4.45}$$
$$\times \delta\left(e_3^{(\alpha)} - e_3^c - e_3^{\bar{c}} - e_3^{cs}\right) J_{n_a}^q\left(z; e_3^c\right) J_{n_b}^{\bar{q}}\left(1 - z; e_3^{\bar{c}}\right) S_{n_a n_b}^+\left(e_3^{cs}\right).$$

As with the factorization theorem in Sec. 4.3.1, we have chosen to write the factorization

---

[8]Here we have slightly extended the usage of the SCET$_+$ nomenclature beyond that which it was originally used in Ref. [105]. In particular, in the case of the signal distribution, there are no global soft modes, and the matching to the effective theory proceeds in quite a different way than for the case of a two prong QCD jet as originally considered in Ref. [105]. Nevertheless, because the effective theory contains a collinear-soft mode, we will refer to it as SCET$_+$.

theorem in terms of $e_2^{(\alpha)}$, $e_3^{(\alpha)}$, and the energy fraction of one of the subjets, z. A brief description of the functions appearing in Eq. (4.45) is as follows:

- $H(Q^2)$ is the hard function describing the production of the on-shell Z bosons in an e^+e^- collision. It also includes the leptonic decay of the Z boson. Following Ref. [85] we assume that the Z boson is unpolarized and so its decay matrix element is flat in the cosine of the boost angle. Non-flat distributions corresponding to some particular decay or production mechanism are straighforward to include.

- $P^{Z \to q\bar{q}}_{n \to n_a, n_b}\left(z; e_2^{(\alpha)}\right)$ describes the decay of the on-shell Z boson into a $q\bar{q}$ pair with momenta along the n_a and n_b axes.

- $J^q_{n_a}\left(z; e_3^c\right)$, $J^{\bar{q}}_{n_b}\left(1 - z; e_3^{\bar{c}}\right)$ are the jet functions describing the collinear radiation associated with the two collinear subjets.

- $S^+_{n_a n_b}\left(e_3^{cs}\right)$ is the collinear-soft function describing the radiation from the $q\bar{q}$ dipole formed by the two collinear subjets.

The basic structure of the factorization theorem, and the radiation described by the different functions, as well as their scalings, are shown schematically in Fig. 4-9. Operator definitions, and one-loop calculations for the operators appearing in the factorization theorem of Eq. (4.45) are given in App. A.5. Because the collinear soft modes are boosted, the collinear soft function does not require a refactorization, as was necessary for the global soft functions, in Sec. 4.3.1.

It is important to emphasize the distinction between our treatment of a boosted Z jet, where we presented a single factorization theorem, and a massive QCD jet, where three distinct factorization theorems were required. While it is obvious that the soft haze region does not exist for a boosted Z jet, the soft subjet region does. However, unlike the case of a massive QCD jet, where the soft subjet region is enhanced by a factor of $1/z_{sj}$ from the eikonal emission factor, no such enhancement exists for the Z decay. Indeed, it was shown in Ref. [85] that the effect of the jet boundary, which would arise from the soft subjet configuration, is power suppressed by $1/Q$. While it would be potentially interesting to

analytically study the jet radius dependence for the signal distribution using the soft subjet factorization theorem, this is beyond the scope of this chapter. We will therefore neglect jet radius effects and write the factorization theorem in Eq. (4.45) with no R dependence.

The factorization theorem of Eq. (4.45) provides an accurate description of the boosted boson signal in the two-prong region of phase space, where $e_3^{(\alpha)} \ll \left(e_2^{(\alpha)}\right)^3$. However, to be able to compare the signal and background distributions, a valid description of the region $e_3^{(\alpha)} \gtrsim \left(e_2^{(\alpha)}\right)^3$ is also required. Unlike for the case of a massive QCD jet, where this region is described by the soft haze factorization theorem, for a boosted Z boson, an accurate description of this region requires matching to the fixed order $Z \to q\bar{q}g$ matrix element. Since the boost of the Z boson is fixed, this corresponds to a hard gluon emission from the $q\bar{q}$ dipole. In the numerical results shown throughout the chapter, we have performed this matching to fixed order, directly within the SCET$_+$ effective theory. The fixed order cross section for $D_2^{(\alpha,\beta)}$ onto which the result of the factorization theorem was matched, was calculated numerically by boosting the leading order $e^+e^- \to q\bar{q}g$ matrix element and performing a Monte Carlo integration. This allows for the consideration of general angular exponents α and β in which case the required integrals are difficult, if not impossible, to evaluate analytically.

4.4 A Factorization Friendly Two-Prong Discriminant

The approach to two-prong discrimination taken in this chapter is to use calculability and factorizability constraints to guide the construction of an observable. Having understood in detail the structure of the $e_2^{(\alpha)}, e_2^{(\beta)}, e_3^{(\alpha)}$ phase space, along with the effective field theories describing each parametric region, we now show how a powerful two-prong discriminant, D_2, emerges from this analysis naturally. After defining the D_2 observable, we discuss some of its interesting properties, and show that the factorization theorems of Sec. 4.3 can be combined to give a factorized description of the observable over the entire phase space.

145

4.4.1 Defining D_2

The goal of boosted boson discrimination is to define observables which distinguish between one- and two-prong jets. As a simplification, we will take the view that both collinear and soft subjets should be treated as two-pronged by the discriminant, while soft haze jets should be treated as one-pronged. Treating both the collinear and soft subjets as two-pronged immediately implies that a marginalization over the soft subjet and collinear subjet factorization theorems will need to be performed to obtain a prediction for the two-prong discriminant. This will be discussed in Sec. 4.4.4. A more sophisticated observable could take advantage of the different fraction of signal and QCD jets in the soft subjet and collinear subjets regions of phase space, and we will give a simple example of such an observable in Sec. 4.5.7.

We will consider discriminants, which we denote $D_2^{(\alpha,\beta)}$, which parametrize a family of contours in the $e_2^{(\beta)}, e_3^{(\alpha)}$ plane, as shown schematically in Fig. 4-10. Such observables can be calculated by marginalizing the double differential cross section [88]

$$\frac{d\sigma}{dD_2^{(\alpha,\beta)}} = \int de_2^{(\beta)} de_3^{(\alpha)} \delta \left(D_2^{(\alpha,\beta)} - D_2^{(\alpha,\beta)}(e_2^{(\beta)}, e_3^{(\alpha)}) \right) \frac{d^2\sigma}{de_2^{(\beta)} de_3^{(\alpha)}} . \qquad (4.46)$$

For the observable $D_2^{(\alpha,\beta)}$ to be calculable using the factorization theorems of Sec. 4.3, the curves over which the marginalization is performed in Eq. (4.46) must lie entirely in a region of phase space in which there is a description in terms of a single effective field theory (up to the marginalization over the collinear and soft subjets). Stated another way, the contours of $D_2^{(\alpha,\beta)}$ must lie either entirely in the one-prong region of phase space, or entirely in the two-prong region of phase space. This condition is also natural from the perspective that $D_2^{(\alpha,\beta)}$ provide good discrimination power, a point which has been emphasized in Refs. [5, 8]. If the contours do not respect the parametric scalings of the phase space, the marginalization cannot be performed within a single effective field theory. A more sophisticated interpolation between the different effective field theories, along the lines of Refs. [10, 34] is then required.

In Sec. 4.2, a power counting analysis was used to show that for $3\alpha/\beta > 2$, the one- and

146
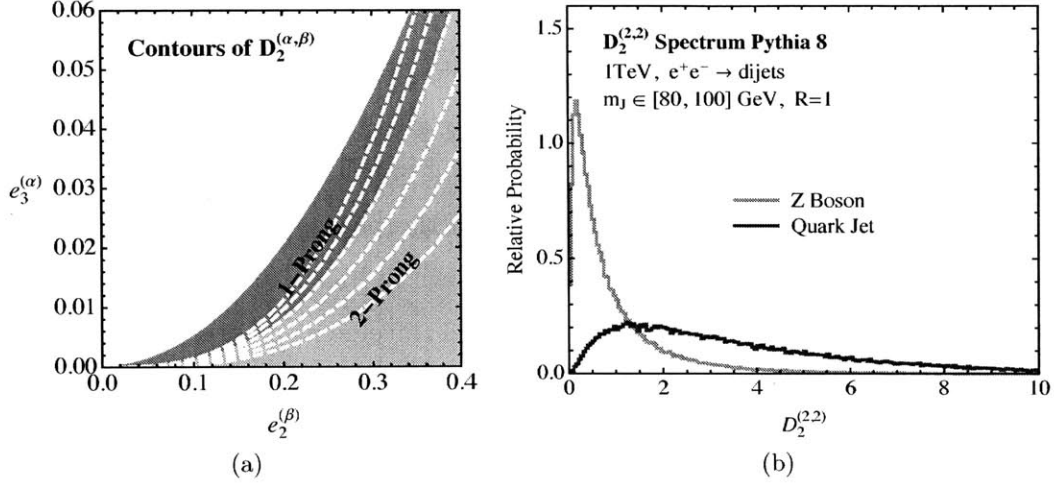
Figure 4-10: a) Contours of the observable D_2 in the $e_2^{(\beta)}, e_3^{(\alpha)}$ plane. b) Sample D_2 spectra for boosted Z bosons and QCD jets, generated in Monte Carlo. Angular exponents $\alpha = \beta = 2$ have been used.

two-prong regions of phase space are parametrically separated, with the contour separating them scaling as $e_3^{(\alpha)} \sim \left(e_2^{(\beta)}\right)^{3\alpha/\beta}$. This implies that, parametrically, the optimal two-prong discriminant formed from $e_2^{(\beta)}$ and $e_3^{(\alpha)}$ is

$$D_2^{(\alpha,\beta)} = \frac{e_3^{(\alpha)}}{(e_2^{(\beta)})^{3\alpha/\beta}} . \tag{4.47}$$

This extends the definition of Ref. [5], which considered the observable $D_2^{(\alpha,\alpha)}$, with equal angular exponents. To simplify our notation, we will often not explicitly write the angular exponents α and β, referring to the observable simply as D_2.

The D_2 observable takes small values for a two-prong jet and large values for a one-prong jet. Its contours in the $e_2^{(\beta)}, e_3^{(\alpha)}$ phase space are shown schematically in Fig. 4-10, along with illustrative Monte Carlo generated spectra for both boosted Z jets and massive QCD jets in e^+e^- collisions. A more detailed discussion of the discrimination power of D_2, as well as the details of the Monte Carlo generation, will be given in Sec. 4.5.

147

4.4.2 Sudakov Safety of D_2

One interesting feature of the D_2 observable is that it is not IRC safe without an explicit cut on $e_2^{(\beta)}$. For every value of D_2, the contour over which the double differential cross section is marginalized passes through the origin of the phase space, where the soft and collinear singularities are located. This feature is shown in Fig. 4-10a. At every fixed order in perturbation theory, this gives rise to an ill-defined (divergent) cross section. However, a resummed calculation of the double differential cross section regularizes the singular region of phase space, and leads to a finite distribution for the D_2 observable. This property is referred to as Sudakov safety [88, 146]. Because Sudakov safe observables are not calculable in fixed order perturbation theory, they do not generically have an α_s expansion, and we will show that the D_2 spectrum exhibits a particularly interesting dependence on α_s.

The regularization of the fixed order singularity in the double differential cross section is achieved by the all orders resummation of logarithmically enhanced terms in the perturbative expansion. In the effective field theory description, this resummation is achieved by renormalization group evolution, and its properties are therefore determined by the form of the SCET anomalous dimensions. To illustrate how the α_s dependence arises from the structure of the renormalization group evolution in SCET, we consider the soft subjet factorization theorem of Sec. 4.3.1 in the leading logarithmic (LL) approximation. The cusp pieces of the anomalous dimensions for the different functions appearing in the factorization are given in Laplace space by (see App. A.3)

$$\mu \frac{d}{d\mu} \log H_{n\bar{n}}^{sj}(z_{sj}, n_{sj}, \mu) = -\frac{\alpha_s C_A}{\pi} \log \left[\frac{\mu^2}{Q^2} z_{sj}^{-2} \right] , \tag{4.48}$$

$$\mu \frac{d}{d\mu} \log J_{n_{sj}} \left(\tilde{e}_3^{(\alpha)} \right) = -2 \frac{\alpha_s C_A}{\pi(1-\alpha)} \log \left[\tilde{e}_3^{(\alpha)} \frac{\mu^\alpha}{Q^\alpha} z_{sj}^{2-\alpha} \right] , \tag{4.49}$$

$$\mu \frac{d}{d\mu} \log S_{n_{sj} \bar{n}_{sj}} \left(\tilde{e}_3^{(\alpha)}; R \right) = \frac{\alpha_s C_A}{\pi(1-\alpha)} \log \left[\tilde{e}_3^{(\alpha)} \frac{\mu}{Q} z_{sj} \right] , \tag{4.50}$$

$$\mu \frac{d}{d\mu} \log S_{n_{sj} n \bar{n}} \left(\tilde{e}_3^{(\alpha)}, B; R \right) = \frac{\alpha_s C_A}{\pi(1-\alpha)} \log \left[\tilde{e}_3^{(\alpha)} \frac{\mu}{Q} z_{sj} \right] , \tag{4.51}$$

where we have used $\tilde{e}_3^{(\alpha)}$ to denote the Laplace conjugate to $e_3^{(\alpha)}$, and we have kept only IR

scales in the logs. Furthermore, we have kept only the terms proportional to C_A so as to resum only the physics associated with the soft subjet. The hard matching coefficient for the soft subjet production is given by the tree level eikonal emission factor

$$H_{n\bar{n}}^{sj(\text{tree})}(z_{sj}, n_{sj}) = \frac{\alpha_s C_F}{\pi z_{sj}} \frac{n \cdot \bar{n}}{n \cdot n_{sj} \, n_{sj} \cdot \bar{n}} \, . \qquad (4.52)$$

Solving the renormalization group equations, and running all functions to the hard scale Q, we then find that in the soft subjet region of phase space the multi-differential cross section can be written to LL accuracy as

$$\frac{d\sigma}{de_3^{(\alpha)} dz_{sj} d\theta_{sj}} = -\frac{\alpha_s^2 C_F C_A}{\alpha \pi^2} \frac{4}{n \cdot n_{sj} \, n_{sj} \cdot \bar{n}} \frac{\log \left[e_3^{(\alpha)} z_{sj}^{-2} \right]}{z_{sj} e_3^{(\alpha)}} e^{-\frac{\alpha_s}{\pi} \frac{C_A}{\alpha} \log^2 \left[e_3^{(\alpha)} z_{sj}^{-2} \right]} \, , \qquad (4.53)$$

exhibiting a familiar Sudakov form.

A complete calculation of the D_2 spectrum requires marginalizing over both the soft subjet and collinear subjet configurations, which we discuss in Sec. 4.4.4. However, to demonstrate the α_s behavior in the simplest manner, we will consider just the soft subjet effective theory. In particular, we will fix the angle of the soft subjet, but allow it to be arbitrarily soft, so as to probe the singular region of phase space. The result is then representative of the contribution from the soft subjet region of phase space. An exactly analogous behavior occurs for the contribution from the collinear subjets region of phase space.

Fixing θ_{sj} to satisfy $n \cdot n_{sj} = 1/2$ (and therefore $\bar{n} \cdot n_{sj} = 3/2$), and restricting to $\alpha = \beta$ for simplicity, the 2-point energy correlation function in the soft subjet region of phase space is simply

$$e_2^{(\alpha)} = z_{sj} \, . \qquad (4.54)$$

The corresponding D_2 distribution is then obtained by marginalizing the multi-differential

cross section of Eq. (4.53)

$$\frac{d\sigma}{dD_2} = \int dz_{sj}\, d\theta_{sj}\, de_3^{(\alpha)}\, \delta\left(D_2 - \frac{e_3^{(\alpha)}}{(e_2^{(\alpha)})^3}\right) \frac{d\sigma}{de_3^{(\alpha)} dz_{sj} d\theta_{sj}} \qquad (4.55)$$

$$\rightarrow \int dz_{sj}\, de_3^{(\alpha)}\, \delta\left(D_2 - \frac{e_3^{(\alpha)}}{z_{sj}^3}\right) \frac{d\sigma}{de_3^{(\alpha)} dz_{sj} d\theta_{sj}},$$

where, in the second line, we have fixed θ_{sj} and so we do not integrate over it. Inserting the multi-differential cross section and fixing θ_{sj}, we then have

$$\frac{d\sigma^{sj}}{dD_2} = -\frac{16}{3}\frac{\alpha_s^2 C_F C_A}{\alpha\pi^2} \int_0^1 dz_{sj} \frac{\log\left[D_2 z_{sj}\right]}{D_2 z_{sj}} e^{-\frac{\alpha_s}{\pi}\frac{C_A}{\alpha}\log^2\left[D_2 z_{sj}\right]} \qquad (4.56)$$

$$= \frac{8}{3}\frac{\alpha_s C_F}{\pi}\frac{e^{-\frac{\alpha_s}{\pi}\frac{C_A}{\alpha}\log^2 D_2}}{D_2},$$

where the sj superscript denotes that this is representative of a contribution from the soft subjet region of phase space. Importantly, because the soft subjet is defined by requirements on IRC safe measurements, the cross section in Eq. (4.56) is a well-defined and in principle measurable quantity.

The α_s dependence in this distribution of D_2 is very surprising. Because D_2 is defined with respect to the 3-point energy correlation function, one would naïvely expect that D_2 only makes sense for a jet with at least three partons. Indeed, if we make an explicit cut on z_{sj}, for example, then D_2 is IRC safe, and first non-zero for a jet with three partons at $\mathcal{O}(\alpha_s^2)$. However, because D_2 without a cut on z_{sj} is not IRC safe, this intuition fails, and in a fascinating way. By resumming the large logarithms of z_{sj} to all orders and then marginalizing, the D_2 distribution calculated in Eq. (4.56) actually starts at $\mathcal{O}(\alpha_s)$! Including emissions to all orders has effectively generated a non-trivial distribution for D_2 at one order *lower* in α_s than when it is first, naïvely, non-zero. Other examples of Sudakov safe observables in the literature have expansions in $\sqrt{\alpha_s}$ [88, 146] or are even independent of α_s [40, 89, 146]. To our knowledge, D_2 is the first example of a Sudakov safe observable for which all-orders resummation reduces the order in α_s when the observable's distribution is

150
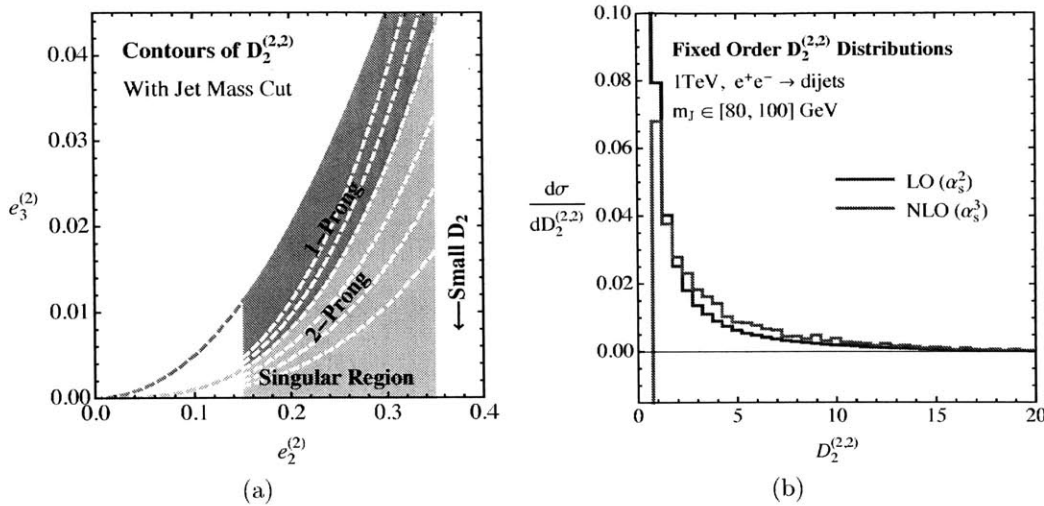
Figure 4-11: a) A schematic depiction of the $e_2^{(2)}, e_3^{(2)}$ phase space in the presence of a mass cut, along with contours of the D_2 observable. b) Leading order (through α_s^2) and next-to-leading order (through α_s^3) distributions for the D_2 observable in the presence of a mass cut as measured on hemisphere jets in e^+e^- collisions.

first non-zero.[9] We re-emphasize that though the distribution of D_2 in Eq. (4.56) is a Taylor series in α_s, it is impossible in purely fixed-order perturbation theory to systematically calculate it.

4.4.3 Fixed-Order D_2 Distributions with a Mass Cut

Although D_2 is not IRC safe without a cut on $e_2^{(\beta)}$, leading to its interesting Sudakov safe behavior, in experimental analyses a jet mass cut will be always be applied. We will therefore be most interested in this case. In Fig. 4-11a we show a schematic depiction of the $e_2^{(\alpha)}, e_3^{(\beta)}$ phase space in the presence of a mass cut for $\alpha = \beta = 2$, along with contours of the D_2 observable. As is indicated in the figure, the mass cut removes the origin of the phase space, making D_2 IRC safe and calculable in fixed-order perturbation theory. It is therefore interesting to study the singularity structure of the fixed-order perturbative expansion of D_2 in the presence of a mass cut.

---

[9]For observables that do not have universal behavior in the ultraviolet [146].

In Fig. 4-11b we show both the leading order (α_s^2) (LO) and the next-to-leading order (α_s^3) (NLO) fixed-order distributions of the $D_2^{(2,2)}$ observable as measured on the most energetic hemisphere jet in $e^+e^- \to$ dijets events at 1 TeV center of mass energy, and with a jet mass cut of $m_J \in [80, 100]$ GeV, in anticipation of our application to boosted Z boson discrimination. However, the detailed range of the mass cut window is irrelevant to the arguments of this section. NLOJet++ [171, 172, 173, 174, 175] was used to generate the distributions. The fixed-order D_2 distribution diverges at small values, and its sign in this region flips order-by-order, characteristic of the Sudakov region. This behavior makes clear the necessity of resummation in the small D_2 region. However, importantly, there is no divergence or other structure at large values of D_2. Instead, the distribution exhibits a tail extending to large values both at LO and NLO, and this behavior is expected to persist to higher orders. This long tail arises from the fact that the upper boundary of the phase space is parametrically far, of distance $\sim 1/e_2^{(\alpha)}$, from the two-prong region of phase space. A schematic depiction of the singularity structure in the $e_2^{(2)}, e_3^{(2)}$ phase space is shown in Fig. 4-11a. The observation that a fixed-order singularity exists only at small values of D_2 is important for the resummation of the observable in the presence of a mass cut. In particular, while resummation in the soft subjet and collinear subjet factorization theorems are necessary to regulate a fixed-order singularity, the soft haze factorization theorem presented in Sec. 4.3.1 is not.

The fixed-order behavior of the D_2 observable is in some ways much more similar to that of a traditional jet or event shape than might naïvely be expected. However, there are some important differences. In particular, a mass cut of $80 < m_J < 100$ GeV has been applied, which is comparable to the location of the Sudakov peak in the mass for a jet of energy 500 GeV. Therefore, unlike in the case of a traditional jet shape, where there is a transition from a region where resummation is important to a far tail region where a fixed order calculation provides an accurate description, in this case, for all values of D_2, there is an overall Sudakov suppression due to the mass cut, in addition to the divergence at small values of D_2. This is however, a small effect in the fixed order distribution compared to the divergence at smaller values, and most importantly, does not require regularization, as it is regulated by the mass

cut.

4.4.4 Merging Factorization Theorems

A complete description of the D_2 observable for background jets requires combining the three factorization theorems presented in Sec. 4.3. This involves both the merging of the soft subjet and collinear subjets factorization theorems, which must be performed before the marginalization over the D_2 contours, as well as the matching between the small D_2 description of the resolved two-prong region and large D_2 description of the unresolved region. We will discuss how the matching is accomplished for these two cases in turn.

Merging Soft and Collinear Subjets

The region of phase space in which two subjets are resolved by the measurement is described by two distinct factorization theorems. These two regions of phase space are separated by the measurement of the two 2-point energy correlation functions, $e_2^{(\alpha)}, e_2^{(\beta)}$. However, in the calculation of D_2, both regions are treated as two-pronged, and the additional 2-point energy correlation function must be marginalized over. Since each effective theory can only be used within its regime of validity, a merged description, valid in both the soft subjets and collinear subjets region of phase space, is required. To accomplish this, we introduce a novel procedure for merging the two factorization theorems.

At a fixed $e_3^{(\alpha)}$, the soft subjet and collinear subjets fill out the $e_2^{(\alpha)}, e_2^{(\beta)}$ phase space, which was shown in Fig. 4-5c. This phase space has also been studied in the context of two angularities measured on a single jet in Refs. [10, 34]. In this case factorization theorems involving only collinear and soft modes exist on the boundaries of phase space, and an additional collinear-soft mode is required in the bulk of phase space. New logarithms exist in the bulk of the phase space, so called k_T logarithms [10], which can either be captured by the additional collinear-soft mode proposed in Ref. [34], or by the interpolation procedure of Ref. [10]. In this case, the factorization theorems involving only the collinear and soft modes do not extend beyond the boundaries of the phase space, and they cannot be directly

153

matched onto one another, as this would neglect the resummation of the k_T logarithms, which are not present in either factorization theorem. We will now argue that the case of interest in this chapter, namely of two resolved subjets, is different. In particular, the soft subjet and collinear subjets factorization theorems extend from the boundaries of phase space, and already contain all the modes required for a description in the bulk of the phase space. In particular no additional modes exist in the bulk region of the phase space. This implies in particular that a description of the entire phase space region can be obtained by a proper merging of the collinear subjets and soft subjet factorization theorems, which is the approach that we will take.

To see that no additional modes are present in the bulk of the phase space, it is sufficient to look for modes which transition between the modes present in the effective theory descriptions in the soft subjet and collinear subjets regions of phase space, and which contribute at leading power. When transitioning from the collinear subjets region of phase space to the soft subjet region of phase space, as is shown schematically in Fig. 4-12a, the collinear modes of one of the jets become the soft subjet and boundary soft modes of the soft subjet factorization theorem. On the other hand, the collinear-soft modes transition to the global soft modes. However, one could possibly be concerned that there exist additional modes which appear as collinear-soft modes on the boundary of phase space where the collinear subjets exist, but which transition to soft subjet modes instead of global soft modes. However, one can immediately see that such modes cannot exist, since the energy fraction of the soft subjet modes is set by the e_2 measurement, while the energy fraction of the collinear-soft modes is set by the e_3 measurement. Since e_3 is fixed, and the transition is occurring only in the $e_2^{(\alpha)}, e_2^{(\beta)}$ phase space, such modes cannot exist. This implies that all contributing modes already exist in either the soft subjet, or collinear subjets factorization theorems. This is a crucial difference from the case of the double differential angularities, which in some sense simplifies the analysis. Since no additional modes exist in the bulk of the phase space, the factorization theorems can be extended from the boundaries, and can be matched onto each other. This will allow for the resummation of all large logarithms. We will now discuss in

154
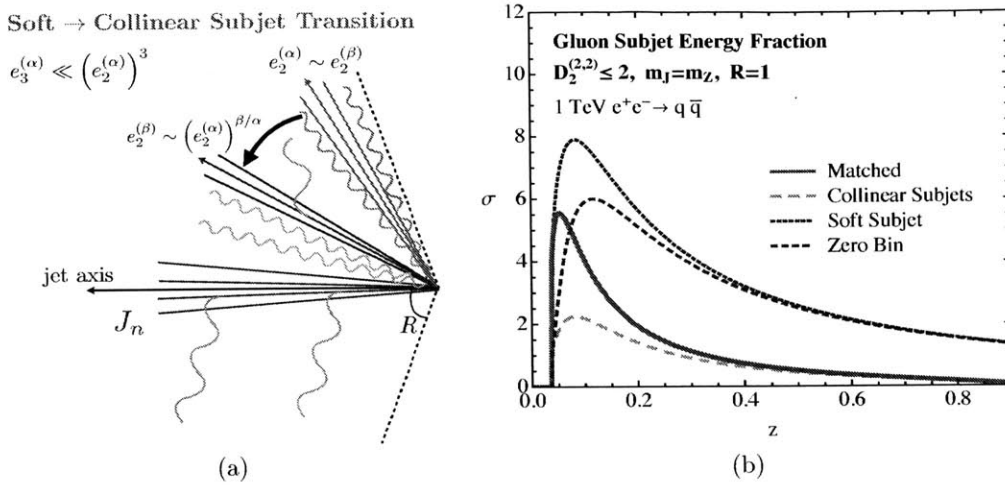
Figure 4-12: a) A schematic depiction of the transition between the soft subjet and collinear subjets regions of phase space. b) Distribution of the energy fraction of the gluon subjet as predicted by the collinear subjets effective theory, the soft subjet effective theory, and the merged description. The collinear zero bin of the soft subjet is also shown.

more detail our implementation of this matching, after which we will see that our argument, presented here based on power counting, for the absence of additional modes, is explicitly realized through our merging procedure.

This suggests then the procedure we will use for interpolating between the collinear subjets and soft subjet factorization theorem, as sketched in Ref. [11], where the soft subjet factorization theorem was originally introduced. It proceeds by implementing a zero bin subtraction [176] in factorization theorem space (the meaning of this will become clear shortly) to remove double counting in the overlapping region between the effective theories. This is a non-trivial and novel example of the zero bin procedure, and demonstrates the general utility of its approach.

Recall that in a standard SCET factorization, the cross section is written as a convolution of a jet function, which describes the collinear physics, and a soft function, which describes the soft physics. To achieve this mode separation without introducing a double counting, the soft limit of the jet function must be subtracted, which is referred to in the literature as a zero bin subtraction. Here we extend this approach to the case of two distinct factorization

155

theorems which describe different regions of a multi-differential phase space, the soft subjet and collinear subjets effective field theories, but which overlap in the bulk of the two-prong phase space. It is important that here we only focus on the two-prong region of phase space; the matching to the one-prong region of phase space will be discussed in Sec. 4.4.4. To perform the matching in the two-prong region of phase space, inspired by the zero-bin procedure, we will write the cross section as a sum of the contributions from the soft subjet factorization theorem and the collinear subjets factorization theorem, with a zero bin contribution to remove the overlap between the effective theories. Explicitly, we write

$$\sigma = \left(\sigma_{sj} - \sigma_{sj}|_{cs}\right) + \sigma_{cs}\,, \tag{4.57}$$

where we have suppressed that at this stage the cross section is still differential in the kinematics of the subjets, so that our notation is not overly cumbersome. The cross section in the soft subjet or collinear subjets regions of phase space are denoted by sj and cs subscripts, respectively. Here the zero bin contribution, which removes the double counting, is given by $\sigma_{sj}|_{cs}$. Explicitly, $\sigma_{sj}|_{cs}$ is obtained by taking the limit of the soft subjet factorization theorem in the power counting of the collinear subjets factorization theorem. The anomalous dimensions and one-loop matrix elements for the collinear zero bin of the soft subjet factorization theorem are given in App. A.4. Each of the three contributions to the cross section given in Eq. (4.57) are associated with their own factorization theorem. However, the contributions to the cross section with the clearest physical interpretation are σ_{cs} and the combined term $(\sigma_{sj} - \sigma_{sj}|_{cs})$, which we will refer to the as the zero bin subtracted soft subjet contribution. It is the contribution which can be interpreted over the entire phase space as the contribution from a soft subjet, and all logarithms contained in this expression are of soft scales.

We specifically subtract the collinear-bin of the soft subjet factorization, and not the soft-bin of the collinear factorization. This is due to the need to cancel the contributions from the boundary soft modes of the soft subjet factorization in the collinear region. Since no analogous mode to the boundary softs is found in the collinear resummation, any soft
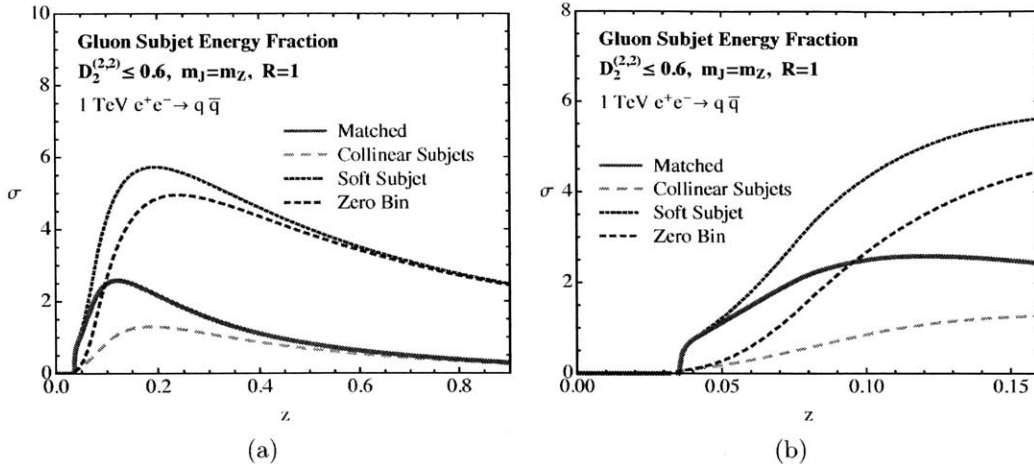
Figure 4-13: a) Distribution of the energy fraction of the gluon subjet as predicted by the collinear subjets effective theory, the soft subjet effective theory, the collinear zero bin, and the matched description. A zoomed version at small z is shown in b).

expansion would miss this contribution, resulting in a logarithm being resummed in an inappropriate *collinear* region of phase space. This is in contrast to what happens when comparing the two subtractions *in the soft region*. So long as one uses the relative transverse momentum of the subjets as the splitting scale of the collinear factorization, the collinear-bin of the soft subjet does match the soft-bin of the collinear factorization in the soft region. This is the result of the merging of various soft scales. In the soft jet collinear-bin, the expanded boundary softs and global soft scales naturally merge, and in the soft-bin of the collinear jets, the global softs and collinear-softs also naturally merge in the soft region. This can be explicitly verified with the canonical scales given in App. A.7. Thus the collinear-bin of the soft subjet is the appropriate subtraction throughout phase space, to remove double counting at all points.

Having defined our merging procedure, implemented through the zero bin, we can now revisit our argument for the absence of additional modes, previously given by power counting, which can be verified from an explicit calculation. Taking the collinear-bin of the soft subjet factorization, and the soft-bin of the collinear subjet factorization, one finds identical fixed order expressions, as well as a one-to-one mapping of the anomolous dimensions between

157

these two re-expanded factorizations. With the merging of the soft scales in the "bins" of the primary factorizations as one enters the soft region then implies they are numerically equivalent. No new logarithms appear in the bulk of phase space, unlike the case of two angularities [10]. This emphasizes that the collinear-soft region is a genuine overlap between the factorizations, with no new structures not already found in the factorizations.

To see visually the effect that this matching has, it is interesting to look at the distribution of the energy fraction of the one of the subjets. In Fig. 4-12b, we plot the distribution of the gluon subjet's energy fraction as computed in the collinear subjets and soft subjet factorization theorems, as well as the energy spectrum for the matched cross section of Eq. (4.57) and zero bin contribution. The energy spectrum is cumulative $D_2 \leq 2$, which is the majority of the two-prong region, and for simplicity we have fixed the jet mass $m_J = m_Z$. The matched contribution smoothly interpolates between the spectrum for the collinear subjets at large values of z, where the collinear subjets factorization theorem is valid, and captures all logarithms of the splitting angle, and that for the soft subjet factorization theorem at small values of z, accurately resumming large logarithms of z. It is also important to note that for large z, the zero bin contribution matches exactly onto the soft subjet contribution, removing its contribution in this region. One can also see that the collinear-bin of the soft subjets cancels the collinear contribution to the soft region, up to power corrections, as argued above. We find that the collinear subjets provides a good description over a large range of values, with the soft subjet factorization theorem only required at small values of z.

In Fig. 4-13a, we show the energy spectra at cumulative $D_2 \leq 0.6$, along with a zoomed version at small values of z, in Fig. 4-13b. This figures makes clear that our matched prediction, computed using our zero-bin approach, reproduces correctly the behavior of the collinear subjets at large values of z, and the soft subjet factorization theorem at small values of z. In particular, in Fig. 4-13b, we see that below $z \sim 0.05$, the soft subjet and matched predictions are indistinguishable.

Although we will not study this case explicitly in this chapter, we have also performed

the matching for gluon jets, where the dominant contribution comes from $g \to gg$ splitting. This case is somewhat interesting due to the fact that the Bose symmetry of the final gluons guarantees that the z distribution is symmetric about $z = 0.5$, leading to peaks in the z distribution due to soft singularities at both $z = 0$ and $z = 1$. Nevertheless, the same matching procedure works identically in this case, and this procedure could therefore also be straightforwardly applied for studying substructure in gluon jets, as would be required for a complete calculation at the LHC.

We have shown here the matched subjet energy spectra for the particular choice of jet radius $R = 1$ at a center of mass energy of 1 TeV for quark jets, as this is the particular case that we will focus on throughout the rest of the chapter. However, we have investigated the properties of the matching away from these parameters. It is important to note that our procedure for merging factorization theorem must be carefully treated at small R. This manifests itself as a breakdown in the zero bin procedure. In particular, for a fixed value of $e_2^{(\alpha)}$, if R is small, then the power counting $e_2^{(\alpha)} \sim z_{sj}$ is invalidated. In other words, for small R there does not exist a region of phase space which contributes to $e_2^{(\alpha)}$ for which z_{sj} is sufficiently small that the soft subjet expansion is valid.

We can bound the specific R that eliminates the soft subjet region by considering the minimum energy fraction accessible to a subjet at a fixed $e_2^{(\alpha)}$:

$$z_{\min} \approx \frac{e_2^{(\alpha)}}{\left(2\sin\frac{R}{2} \right)^\alpha} \, . \tag{4.58}$$

As a necessary condition for a soft subjet, one must fulfill the condition:

$$z_{\min} \sim e_2^{(\alpha)} \to 1 \sim \left(2\sin\frac{R}{2} \right)^\alpha \, , \tag{4.59}$$

and so $R \sim 1$ for the soft subjet to contribute. To eliminate the soft subjet then requires $R \ll 1$ and to still have valid collinear subjet regions requires that R and $e_2^{(\alpha)}$ are related as:

$$1 \gg R^\alpha \gg e_2^{(\alpha)} \, . \tag{4.60}$$

Finally, one should distinguish a fixed mass jet from a fixed $e_2^{(\alpha)}$. In the case $\alpha = 2$, since $e_2^{(2)} = \frac{m_J^2}{E_J^2}$, by varying E_J or R, we can open or close the soft subjet region.

This appears in the zero bin by the fact that the zero bin subtraction is greater in all regions than the soft subjet, leading to a negative total cross section. We find numerically that this occurs for $R < 0.5$ for the case of $m_J = 90$ GeV, and $Q = 1$ TeV. This value depends fairly sensitively on m_J and Q, or equivalently $e_2^{(\alpha)}$. In this case, only the collinear subjets factorization theorem should be used, and it is valid throughout the entire available phase space. In this chapter we focus primarily on the case of fat jets, defined with $R = 1$, and therefore it is necessary to perform the matching between the soft subjet region and the collinear subjets region for jets of energy 500 GeV. However, in Sec. 4.5.3, we perform a brief survey of different R values, comparing our analytic predictions with distributions from Monte Carlo generators. A more phenomenological study of the importance of the matching for different physics processes of interest for an e^+e^- collider, the LHC, or even a possible 100 TeV collider, where even higher boosts can be achieved, would be interesting, but is well beyond the scope of our initial investigation and can be straightforwardly treated using our techniques.

While we have used a zero bin procedure to perform the matching between the collinear subjets and soft subjet factorization theorems, it is also possible to develop a dedicated effective field theory valid when the soft subjet becomes collinear. This effective field theory is related to our zero bin contribution, and has been developed in [36]. While we believe that this approach is nice in principle, for the observable D_2, we find that such an effective field theory has a vanishing region of validity, as can be seen from the zero bin contribution in Fig. 4-12b, and Figs. 4-13a and 4-13b. We therefore believe that our use of the zero bin, as generalized to distinct factorization theorems, represents a natural approach to the merging of the distinct factorization theorems. However, we acknowledge that this is an observable dependent statement, and there may be cases where there is a sufficiently large region of overlap between the soft subjet and collinear subjets effective theories, and in this case it might prove useful to have a separate effective field theory description which is valid in the

case that the soft subjet becomes collinear.

Matching Resolved to Unresolved Subjets

An important feature of the D_2 observable is that its contours respect the parametric scaling of the phase space, as emphasized in Fig. 4-10. This implies that the marginalization over the contours defining the observable can be performed at small D_2 entirely within the merged effective theory of Sec. 4.4.4, and at large D_2 within the soft haze effective field theory. Hence the matching between these two different descriptions can be performed at the level of the D_2 distribution instead of at the level of the double differential cross section, which is a great simplification, and primary feature of the D_2 observable.

The soft haze factorization theorem presented in Sec. 4.3.1 first contributes to the shape of the D_2 distribution at two emissions, the first order at which $e_3^{(\alpha)}$ can be non-zero (technically at next-to-next-to-leading logarithmic prime order, NNLL′, in the logarithmic counting). Since our focus is on an initial investigation of the factorization properties of two-prong discriminants, the necessary two-loop calculation is beyond the scope of this chapter. Naïvely, this implies that since the merged effective field theory describing the two-prong region of phase space is only valid for $D_2 \lesssim 1$, our predictions should not be extended beyond $D_2 \lesssim 1$. However, we will argue that because of the structure of fixed order singularities for the D_2 observable, extending our two-prong factorization theorems to large D_2 will provide an accurate description of the D_2 distribution for a wide range of E_J and R.

As shown in Sec. 4.4.3, there does not exist a fixed order singularity at large D_2. In particular, this implies that if extended into this region, the factorization theorems valid at small D_2 will not diverge. Furthermore, one in fact expects that they provide a reasonable description of the shape. They contain both an overall Sudakov factor for the $e_2^{(\beta)}$ scale of the jet, and also provide a description of the internal structure of the jet in terms of splitting functions (in the case of the collinear subjets factorization). While the splitting function does not exactly reproduce the matrix elements in the soft haze factorization theorem, it provides a good description of them. We believe that this is a consistent approach which

suffices for this initial investigation.

Perhaps the most important fixed order correction not captured in the subjet factorition for D_2 is simply the endpoint of the distribution, which arises from the kinematic boundaries of the phase space. Since we will normalize our distributions to 1, in order to compare to the Monte Carlo generators, the height of the peak is correlated with the endpoint. Matching to the soft haze region would give the resummed distribution the correct endpoint in the tail, and thus can shift the peak up in general. This endpoint is sensitive to the specific R and E_J of the jet, as well as to the values of the angular exponents α and β. Recall that since the Monte Carlo generators respect momentum conservation, they always terminate their distributions before the physical endpoint of the spectrum. We will also see how this disagreement in the tail region changes as a function of R and E_J in Secs. 4.5.3 and 4.5.4 respectively. However, for the case of dijets produced at a center of mass energy of 1 TeV, with a jet mass cut of $80 < m_J < 100$ GeV, as is relevant for boosted boson discrimination, and on which we primarily focus throughout this chapter, we will see that this discrepancy in the tail region is minimal, and we will find good agreement between our analytic calculations and the Monte Carlo predictions. It would of course be interesting to perform the complete two-loop calculation in the soft haze region of phase space; however, we believe that this would have a minor effect for a substantial range of parameter space. Nevertheless, the proper inclusion of this region of phase space would also be interesting from a resummation perspective, as it would require matching between two distinct factorization theorems involving a different number of resolved jets, instead of the more familiar case of matching a resummed distribution to a fixed order calculation. We leave further investigations of this to future work.

4.5 Numerical Results and Comparison with Monte Carlo

We now present numerical results for signal and background distributions for the D_2 observable in e^+e^- collisions. We give a detailed comparison with Monte Carlo, at parton level in Secs. 4.5.1 through 4.5.4 and including hadronization in Sec. 4.5.5. We then study the discrimination power of D_2 analytically in Sec. 4.5.6, and comment on the optimal choice of

angular exponents. In Sec. 4.5.7 possible observables which go beyond D_2, and separately resolve the soft subjet, and collinear subjets region of phase space, and how these could be used for possible improvements to boosted boson discrimination.

Throughout this section we use FASTJET 3.1.2 [100] and the `EnergyCorrelator` FAST-JET CONTRIB [100, 101] for jet clustering and analysis. All jets are clustered using the e^+e^- anti-k_T metric [94, 100] using the WTA recombination scheme [33, 89], with an energy metric.[10]

4.5.1 Comparison with Parton-Level Monte Carlo

Previous studies of boosted boson discrimination with ratios of IRC safe jet observables have relied entirely on Monte Carlo simulations. While the implementation of both the perturbative shower and hadronization are well-tuned to describe simple event-wide observables, jet substructure observables probe significantly more detailed correlations. For the particular case of observables sensitive to two-prong structure, their discrimination power is sensitive to the description of massive QCD jets in the phase space region where the jets are dominated by a resolved splitting. One might naïvely expect that this region of phase space is sensitive to the implementation of the parton shower model, and we will see that this is indeed the case.

While a comparison to recent LHC data on jet substructure observables (for example: [56, 57, 59, 74, 70, 126]) is possible, the lack of analytic calculations means that it is difficult to disentangle perturbative from non-perturbative effects. In this section we compare the results of our analytic calculation for D_2 with a number of Monte Carlo generators at parton level, focusing in particular on the small D_2 region.[11] This allows for a detailed probe of the simulation of two-prong jets in QCD by the perturbative shower (for a discussion of some

---

[10]We thank Jesse Thaler for use of a preliminary version of his code for WTA in e^+e^- collisions. This code is now available in the FASTJET CONTRIB.

[11]One should always be wary of comparisons of Monte Carlo generators at parton level which employ different hadronization models. Our comparisons at parton level presented in this section are to set the stage for fully hadronized comparisons in the following section. However, we take the view that a parton shower should achieve, to the greatest extent possible, a clean separation between perturbative and non-perturbative physics, and therefore should provide an accurate description of observables both at parton and hadron level.

other variables, see Ref. [177, 178]). A large number of implementations of the perturbative shower exist, and are implemented in popular Monte Carlo generators (for reviews, see e.g. [179, 180, 181, 182, 183]). Some examples include PYTHIA [108, 109], a p_T-ordered dipole shower; VINCIA [149, 150, 151, 152, 153, 154], SHERPA [184, 185], ARIADNE [186], and DIRE [187], dipole-antenna showers; and HERWIG++ [110, 111, 112, 113], an angular-ordered dipole shower.[12]

As representative of these different Monte Carlo shower implementations, we will use the following Monte Carlo generators throughout this section:

- PYTHIA 8.205

- VINCIA 1.2.01 with a p_T-ordered shower

- VINCIA 1.2.01 with a virtuality-ordered shower

- HERWIG++ 2.7.1

All Monte Carlos were showered with default settings except for the caveats listed below and requiring two-loop running of α_s with $\alpha_s(m_Z) = 0.118$. The different shower evolution variables within the VINCIA Monte Carlo enables a study of their effects. For background distributions, we generate $e^+e^- \to$ dijets at 1 TeV center of mass energy and study the highest energy $R = 1.0$ anti-k_T jet in the event. For signal distributions in PYTHIA and VINCIA, we generate $e^+e^- \to ZZ$ events with both Zs decaying hadronically. For HERWIG, the fixed-order signal distributions are generated in MADGRAPH5 2.1.2 [107] and showered in HERWIG. All jets are required to have a mass in the window $m_J \in [80, 100]$ GeV. In all plots shown in this section, hadronization has been turned off in all Monte Carlos. Fixed-order matching was also turned off in VINCIA.

Fig. 4-14 compares our analytic prediction for the $D_2^{(2,2)}$ spectrum to the parton-level Monte Carlo simulations in both background (Fig. 4-14a) and signal (Fig. 4-14b) samples. The details of the scale variations used to make the uncertainty bands will be explained in

---

[12]HERWIG++ also has the option for a dipole-antenna shower implementation [188] though we will not use it here.
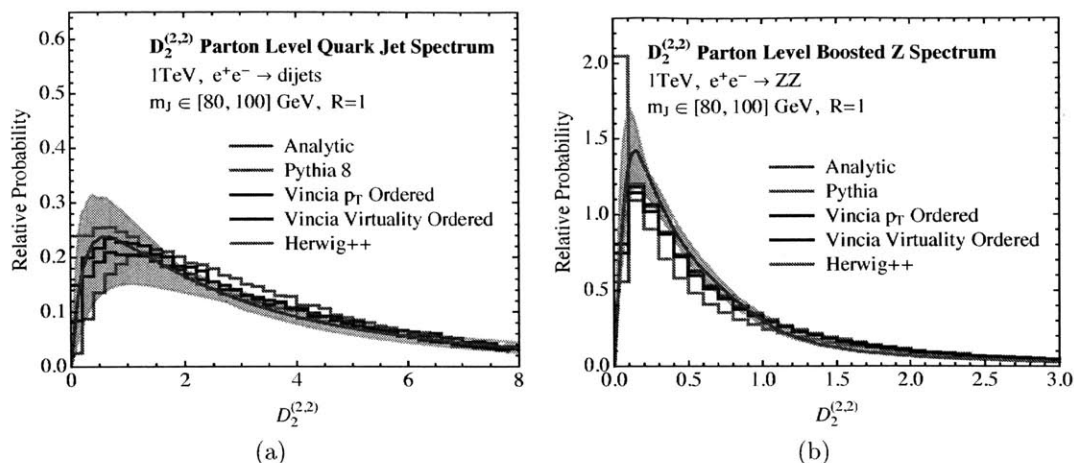
Figure 4-14: A comparison of our analytic prediction for $D_2^{(2,2)}$ compared with the parton-level predictions of the PYTHIA, VINCIA and HERWIG Monte Carlos. a) The D_2 distributions as measured on QCD background jets. b) The D_2 distributions as measured on boosted Z boson signal jets. The solid line is the central value of our analytic calculation and the shaded bands are representative of perturbative scale variations.

Sec. 4.5.2, but the pinch in the uncertainties should not be taken as physical. All Monte Carlos have similar distributions as measured on signal jets, though HERWIG is more peaked at small values than the other generators. Our analytic prediction, shown with perturbative scale variation, agrees well with the Monte Carlo generators. On background jets, however, the distributions are distinct, especially at small values of D_2. Small D_2 is the region where the jet has a two-prong structure, but unlike for signal jets, for background jets that structure is not generated by a hard matrix element. In the case of collinear subjets, it is generated by a hard splitting function, while for a soft subjet, it is generated by an eikonal emission. Because it is an antenna shower, and therefore accurately describes the emission of wide-angle, soft gluons from leading-color dipoles, the VINCIA distribution agrees the best with our calculation. Also, p_T-ordering agrees better than virtuality-ordering, as expected because the scales at which the functions in the factorization theorem for D_2 are evaluated are essentially the relative p_T of the dominant emissions in the jet. This agrees with the conclusions of Ref. [189, 153], where it was shown that for antenna showers, the choice of p_T as a shower ordering variable absorbed all logarithms to $\mathcal{O}(\alpha_s^2)$. The PYTHIA distribution

165

is shifted right as compared to VINCIA and our calculation, which appears to be due to the way that PYTHIA populates the soft, wide angle region of phase space.[13] HERWIG, though a dipole shower like PYTHIA, agrees well with our calculation as shown in Fig. 4-14a.

For reference, in App. A.9 we show a collection of $e_2^{(2)}$ distributions at both parton and hadron level for each of the different Monte Carlo generators. Since $e_2^{(2)}$, which is related to the jet mass by Eq. (4.8), is set by a single emission, the agreement between the different generators, particularly at parton level, is significantly better than for the D_2 observable. This further emphasizes the fact that the D_2 observable offers a more differential probe of the perturbative shower, going beyond the one emission observables on which Monte Carlo generators have primarily been tuned.

In the following sections we will study the partonic D_2 distributions in more detail. We will restrict ourselves to comparing and contrasting p_T-ordered VINCIA and PYTHIA for a few reasons. First, as exhibited in Fig. 4-14a, these Monte Carlos represent the largest spread in their predicted D_2 spectra. HERWIG, while it performs very similarly to VINCIA, has a different hadronization model than PYTHIA and VINCIA. So, directly comparing PYTHIA and VINCIA minimizes any implicit hadronization effects when comparing the Monte Carlos at parton level. There are still differences due to the cutoff of the perturbative shower, which will be discussed in Sec. 4.5.2.

4.5.2   Monte Carlos and Perturbative Scale Variation

The fact that, in particular, the p_T-ordered VINCIA distribution for D_2 as measured on background agreed with our calculation while the PYTHIA distribution disagreed in the small D_2 region can be understood and quantified further. The bulk of the disagreement between our analytic calculation and PYTHIA, illustrated in Fig. 4-14a, occurs near the peak of the D_2 distribution. It is well-known that for many observables perturbative uncertainties tend to be significant in the peak region of the distribution. Therefore, it is possible that the difference between the p_T-ordered VINCIA and PYTHIA D_2 distributions can fully be explained by large

---

[13]We have checked that this conclusion remains true with and without including matrix element corrections in the parton shower. We find little difference between the two cases.

perturbative uncertainties. In this section, we will provide evidence showing that this is not the case and that this discrepancy arises from the different way that the two Monte Carlos treat the soft subjet region of phase space.

To estimate perturbative uncertainties in our resummed analytic calculation, the standard procedure is to vary the scales that appear in the calculation by factors of 2. This is at the very least a proxy for the sensitivity of the cross section on these scales. Because our factorization theorems contain many functions, as well as merging of distinct factorization theorems, in principle there are numerous scales that could be varied, a complete analysis of which is beyond the scope of this chapter. A complete list of the variations considered as well as the resummation procedure can be found in App. A.7, while here we only summarize. In all factorizations theorems, we vary the subjet splitting scales, the in-jet soft radiation scales, the out-of-jet soft radiation scales, as well as where the freeze-out for the Landau pole occurs in the running of α_s. We do not separately vary the scale in the soft subjet factorization theorem and the collinear zero bin to ensure that the zero bin subtraction is implemented correctly. The scale variation band for the total cross section is then taken as the combined band for all possible combinations of these scale variations. The soft subjet cross section displays a particular sensitivity to the out-of-jet scale setting, since the running between the boundary soft modes and the out-of-jet modes forces the soft subjet energy spectrum to vanish at the jet boundary,[14] though the fixed order cross section probes the soft divergence in this region. Thus we also consider several different schemes for handling the out-of-jet scale setting. We believe that our scale variation bands are representative, and this is supported by the agreement with the Monte Carlo.

Having understood the perturbative uncertainty bands, we now discuss in more detail the discrepancy between the different Monte Carlo generators arising at small values of D_2, as exemplified by the difference between the p_T-ordered VINCIA and PYTHIA distributions. To understand the origin of this discrepancy, we begin by understanding the effect of the soft subjet region of phase space in our analytic calculation. This is possible due to our complete separation of the phase space using the energy correlation functions. In particular, because

---

[14]As explained in Ref. [11], this is connected with the buffer region of Ref. [190].
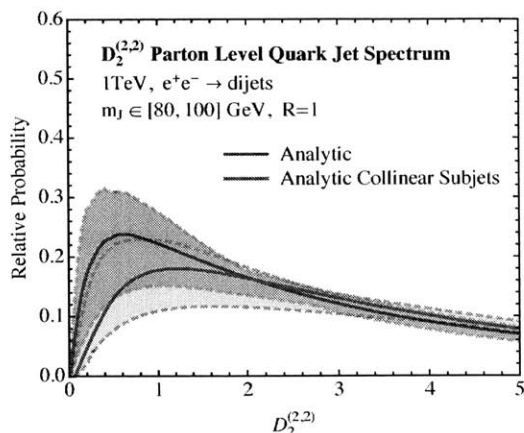
Figure 4-15: Analytic prediction for the D_2 distribution for background QCD jets including the envelope of the perturbative scale variation, as compared with an analytic calculation including just the collinear subjets region of phase space. The effect of the soft subjet region of phase space is clearly visible at small values of D_2.

we have formulations of distinct factorization theorems in the soft subjet and collinear subjets regions of phase space, we can make an analytic prediction for the contribution arising just from the collinear subjets region of phase space. In Fig. 4-15 we show a comparison of the D_2 distribution for background QCD jets as computed using our complete factorization theorem, incorporating both the soft subjet and collinear subjets region of phase space, as compared with the calculation incorporating only the collinear subjets region of phase space. Comparing the two curves, we are able to understand the effect of the soft subjet region of phase space. In particular, we see that the soft subjet has a considerable effect on the distribution at small values of D_2, giving rise to a more peaked distribution, with the peak at smaller values of D_2, as compared to the result computed using only the collinear subjets region of phase space. Although the perturbative error bands are large, the systematic effect of the soft subjet region of phase space is clear.

One further feature of the D_2 distributions, which is made clear by Fig. 4-15, is that the full D_2 distribution is not the result of a single Sudakov peak, and therefore our intuition about the behavior of different orders in the perturbative expansion, and the behavior of scale variations from traditional event shapes fails. In particular, while it is generically the

168

case for traditional event shape distributions that lower order resummed results overshoot in the peak region, it is not at all clear that this behavior should be true for D_2, and indeed it is not observed. Instead, the contribution from the collinear subjets alone is expected to undershoot the peak of the D_2 distribution, since it does not incorporate the soft subjet region of phase space. The final contribution is then obtained as a superposition of two distinct Sudakov peaks, and can therefore behave quite differently from traditional event shapes.

Monte Carlo descriptions of the perturbative shower should provide a similar description of collinear physics, but can differ in their description of soft wide angle radiation. Some of these differences were discussed in Sec. 4.5.1. As discussed earlier, because VINCIA is a dipole-antenna shower, it should accurately describe both the hard collinear and soft wide-angle regions of phase space. Because small values of D_2 are sensitive to both collinear and soft physics, the fact that the PYTHIA distribution at small D_2 is distinct suggests that its description of soft wide-angle physics is the reason.[15] The difference observed in our analytic calculation arising from the soft subjet region of phase space is similar to that observed between the p_T-ordered VINCIA and PYTHIA Monte Carlo distributions. It is therefore interesting to investigate whether the difference in Monte Carlo distributions can arise from different descriptions of wide angle soft radiation. We will now provide evidence that this is indeed the case. It is important to emphasize, however, that since we perform this comparison at parton level, there is some ambiguity in effects due to the perturbative cutoff of the shower, and those arising from different descriptions of wide angle soft radiation. In particular, the Monte Carlos will in general have different low-scale p_T cutoffs at which the perturbative parton shower is terminated. Varying this scale can potentially greatly increase or decrease the number of soft emissions because the value of α_s in this region is large. In particular, for the versions of PYTHIA and VINCIA that were use to generate events

---

[15]Part of the reason for why PYTHIA seems to not correctly describe the soft, wide-angle region of phase space may be due to the fact that while it uses kinematics of dipoles in its shower, it still uses the Altarelli-Parisi splitting functions as an approximation of the squared matrix element. The dipole and its emission is then boosted to the appropriate frame, which may over-populate the soft wide-angle region of phase space as compared to the eikonal matrix element. We thank Torbjörn Sjöstrand and Peter Skands for detailed discussions of this point.
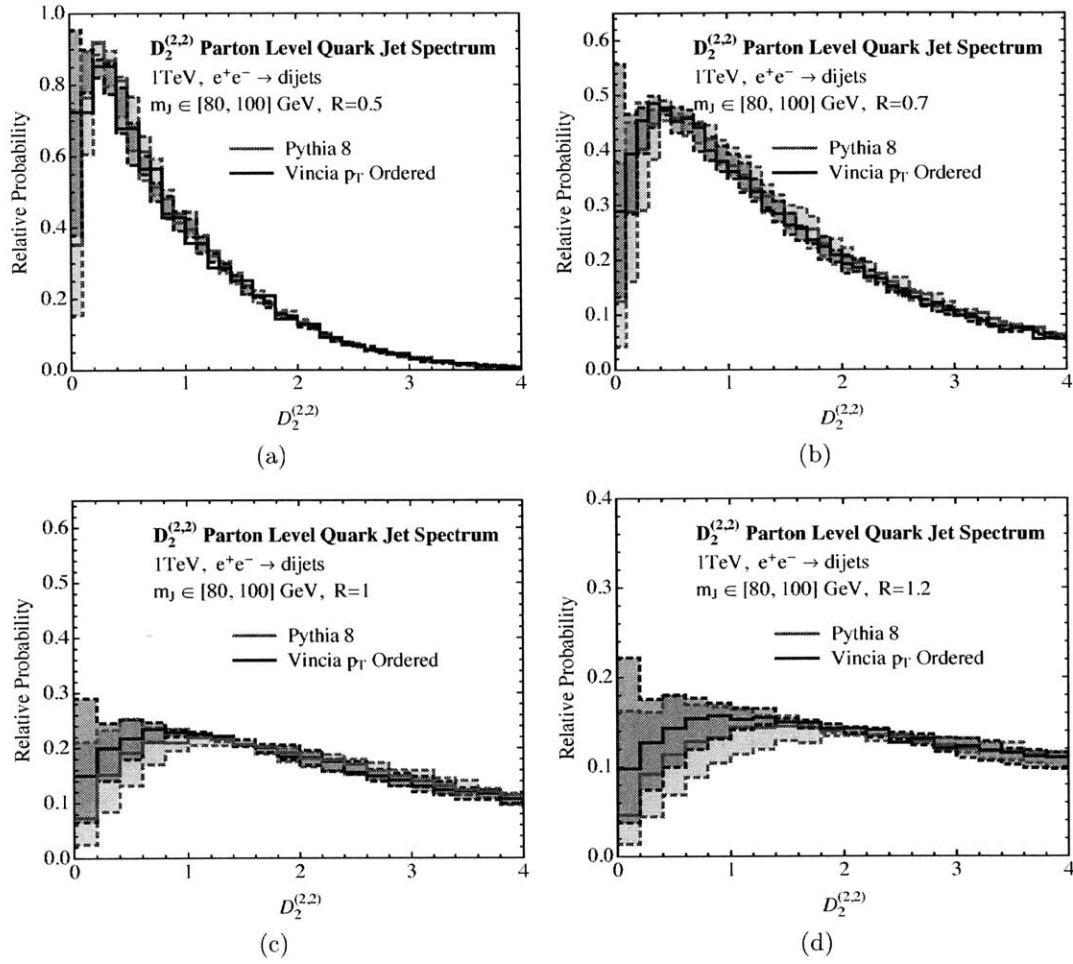
Figure 4-16: Comparison of the D_2 distribution for background QCD jets in PYTHIA and p_T-ordered VINCIA for different jet radii. In each plot, the central value is obtained using a shower cutoff of 0.8 GeV, and the uncertainty bands are generated by varying this cutoff between 0.4 GeV, and 1.2 GeV.

in Fig. 4-14, the cutoff in PYTHIA is 0.4 GeV, while the cutoff in VINCIA is 0.8 GeV. Indeed, these are the default values for these showers. Therefore, we expect that the PYTHIA parton shower produces more soft emissions than VINCIA, which would increase the value of D_2, and potentially also contribute to the observed difference.

To attempt to disentangle the effects of the shower cutoff from differences in the modeling of soft radiation, in Fig. 4-16, we consider Monte Carlo predictions of the D_2 distribution as measured on QCD jets, with different jet radii, namely $R = 0.5, 0.7, 1.0, 1.2$. By using different jet radii, we can control the importance of the soft subjet region of phase space. With small jet radii, the soft subjet region of phase space does not exist, while it becomes increasingly important as the jet radius is increased. Analytic predictions for the D_2 distribution for different values of the jet radius, R, will be given in Sec. 4.5.3. Here we compare the p_T-ordered VINCIA and PYTHIA Monte Carlo. To generate the central values of the curves, we have used a cutoff of 0.8 GeV, and the uncertainty bands are generated by varying this cutoff from 0.4 GeV to 1.2 GeV, to understand its effect. From Fig. 4-16 we see that while there is a relatively large uncertainty band from varying the perturbative cutoff of the shower, the uncertainty bands for the two generators become systematically more offset as the jet radius is increased. At $R = 0.5$ or 0.7, there is very small distinction, while for $R = 1.0$ or 1.2, there is a relatively large offset. This suggests that while some of the distinction seen between the Monte Carlo distributions in Fig. 4-14 is due to the different treatment of the cutoff of the perturbative shower, there is a difference arising from the different description of perturbative wide angle soft radiation. The dipole-antenna based shower, VINCIA, which is expected to provide an accurate description of wide angle soft radiation seems to agree best with our analytic calculation, suggesting that PYTHIA is not providing a complete description of wide angle soft radiation, which can play an important role in jet substructure studies.

This analysis also shows some of the difficulties in disentangling perturbative from non-peturbative effects, and the importance of having analytic calculations and precise theoretical control of different phase space regions to do so. However, by measuring sufficiently many observables on a jet, we are able to isolate distinct phase space regions and study in detail the

171

extent to which Monte Carlo parton showers reproduce the physics in the different regions. D_2, or similar jet substructure observables, could therefore be powerful tools for tuning Monte Carlos, both to formally-accurate perturbative calculations, as well as data. In the remaining sections of the chapter, we will use the default shower cutoffs in the Monte Carlo generators, as was done in Fig. 4-14, and will not show uncertainty bands on our Monte Carlo distributions from varying this parameter.

4.5.3   Analytic Jet Radius Dependence

As demonstrated in the previous section, the region of small D_2 is a sensitive probe of the dominant soft or collinear structure in the jet. It is therefore interesting to study the jet radius dependence of D_2 analytically, because the relative size of soft subjet and collinear subjets contributions to D_2 will depend on the jet radius. At large jet radius, as shown earlier, the soft subjet region is an important contribution at small D_2, but as the jet radius decreases, the collinear subjets should dominate. In this section, we will study the jet radius dependence of D_2 and compare our analytic calculation to Monte Carlo. This will also demonstrate that our analytic calculation accurately describes the R dependence of the D_2 distribution. As in the previous section, we will restrict this study to p_T-ordered VINCIA and PYTHIA showers, and will take the jet radius to be $R = 0.5, 0.7, 1.0, 1.2$, which are representative of a wide range of values of experimental interest. Larger values of R can be straightforwardly studied with our approach, but are of less phenomenological interest. It is expected that for smaller values of R logarithms of R may become numerically important [134, 191, 192], so we do not consider them here.

Comparisons of parton level Monte Carlo results from both p_T-ordered VINCIA and PYTHIA to our analytic calculation are shown in Fig. 4-17. Since we scan over a range of jet radii, perturbative uncertainties for each R value are not as extensively explored as earlier with $R = 1$, and are only meant as a rough estimate of the perturbative uncertainty. Our focus here is simply to show that the scaling behavior with R between our analytic calculation and the Monte Carlos agree. There is excellent agreement between the Monte

172

Carlo results and our analytic calculations over the entire range of R values, with p_T-ordered VINCIA providing a more accurate description of the analytic calculation than PYTHIA. The size of the region over which there is a large disagreement between the Monte Carlos increases monotonically with R, and for $R \gtrsim 1$, there is even considerable disagreement in the position of the peak of the distribution between the generators. However, for smaller values of R, the discrepancy between VINCIA and PYTHIA at small D_2 is reduced. Indeed the position of the peak of the distribution coincides between the different generators for $R = 0.5$, with only a small discrepancy in the bin at lowest D_2. However, here hadronization will play an important role, smearing out this effect. The effect of hadronization, and its implementation in our analytic calculation, will be discussed in Sec. 4.5.5.

For jet radii of $R = 0.7, 1.0, 1.2$ our analytic calculation consists of both collinear subjets and soft subjet contributions. For $R = 0.5$, however, we only include the contribution from collinear subjets, which is guided by our matching procedure between the collinear subjets and soft subjet factorization theorems, as discussed in Sec. 4.4.4. For a fixed jet mass, as the value of R is decreased, the region of validity of the soft subjet factorization theorem vanishes rapidly. For jet masses in the range $80 < m_J < 100$ GeV, and $Q = 1$ TeV, we find that between $R = 0.7$ and $R = 0.5$ the region of validity of the soft subjet rapidly shrinks to zero, and there should not be a transition between the collinear subjets factorization theorem and the soft subjet factorization theorem. Because of this, for the value of $R = 0.7$, our perturbative error bands are more extensive, and are taken as the envelope of curves both that include the matched soft jet, and curves that do not. While this is certainly over conservative in the error estimate, we have included this to emphasize this point. This feature is also seen explicitly in the plots of Fig. 4-17, where the region of disagreement between the different Monte Carlo generators is squeezed towards zero. A similar effect occurs as the energy (or p_T) of the jet is increased with a fixed jet mass, which will be discussed in Sec. 4.5.4.

As discussed in Sec. 4.5.2, this observation of the R dependence provides support that the discrepancy between the Monte Carlo generators, as represented by PYTHIA and VINCIA,

173
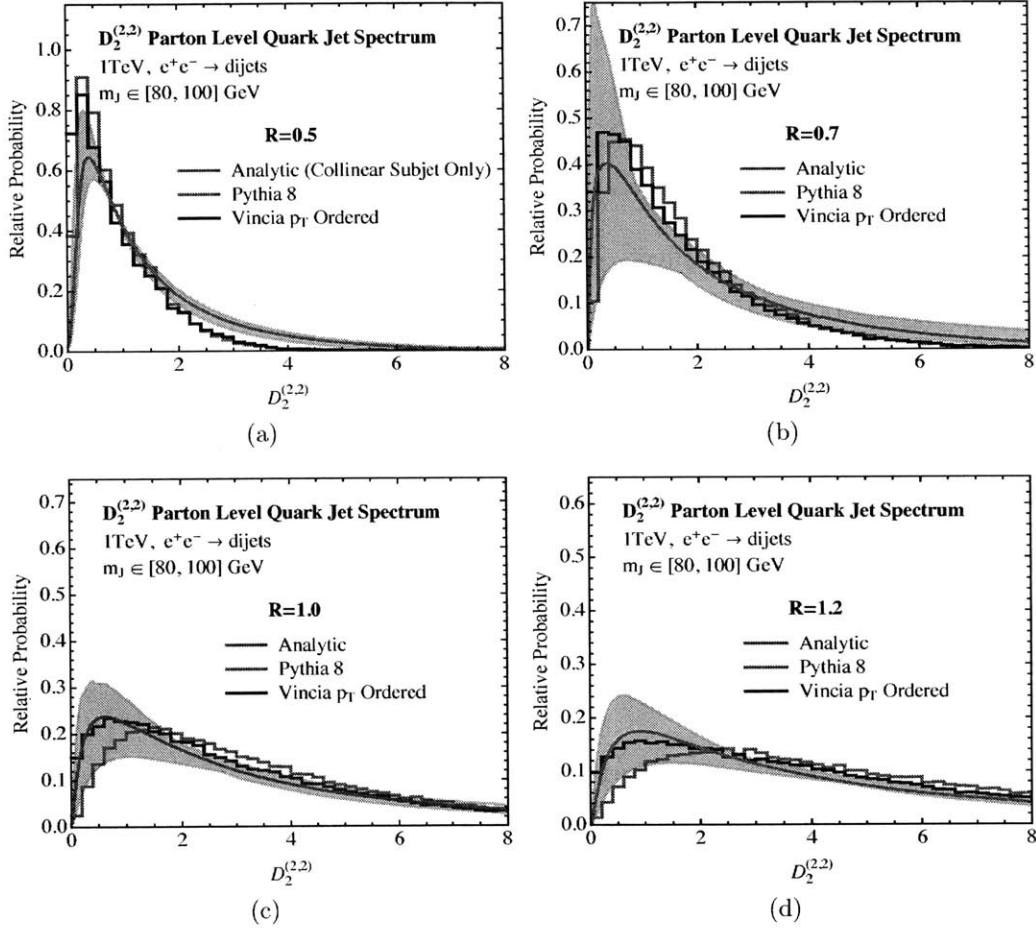
Figure 4-17: Comparison of QCD background D_2 distributions from p_T-ordered VINCIA and PYTHIA to our analytic prediction as a function of the jet radius, R. The values $R = 0.5, 0.7, 1.0, 1.2$ are shown in Figures a)-d), respectively. In the analytic prediction for $R = 0.5$, only the collinear subjets factorization theorem is used, while for all other values of the jet radius the analytic calculation includes contributions from both the collinear subjets and soft subjet factorization theorems.

174

is dominantly due to PYTHIA's description of wide angle soft radiation. The disagreements between Monte Carlo generators is thus highly sensitive to the degree to which soft subjets can impact the distributions, a feature which should be taken into account when performing jet substructure studies. Throughout the remainder of this chapter, we will study the case $R = 1$ exclusively, because both collinear subjets and soft subjet regions of phase space must be included and that radius is relevant to a large number of jet substructure studies using fat jets.

4.5.4 Analytic Jet Energy Dependence

In addition to studying the dependence on the jet radius as a probe of the importance of the soft subjet and of the Monte Carlo description of the shower, it is also interesting to study the dependence of the D_2 distribution on the energy of the jet, with a fixed mass cut. For highly energetic jets, one expects that the soft subjet will play a negligible role, as the region of validity of the soft subjet factorization theorem shrinks as the energy of the jet is increased, as long as the mass of the jet is kept fixed. On the other hand, since we keep the jet radius used in the clustering fixed, the angular separation of the collinear particles decreases with energy, but the phase space for wide angle global soft radiation increases considerably. This radiation is present both in the collinear subjets and soft haze factorization theorems. It is also of course present in the soft subjet factorization theorem, although we have argued that we expect this to give a small contribution. Studying the jet energy dependence therefore probes the behavior of the generators in a fashion complementary to the R dependence.

In this section, we study the perturbative D_2 distribution for center of mass energies ranging from 500 GeV to 2 TeV, for a fixed jet radius of $R = 1$, and with a fixed mass cut of $80 < m_J < 100$ GeV. This region of energies covers the majority of the phenomenologically interesting phase space available at the LHC. We will also perform a more detailed study at LEP energies in Sec. 4.6. For our resummation, we require (amongst other things), that $e_2^{(\alpha)} \ll 1$. For the case of $\alpha = 2$ for which we will be most interested, this corresponds to the assumption $e_2^{(2)} = m_J^2/E_J^2 \ll 1$. For a mass cut around the Z pole mass, this expansion

175
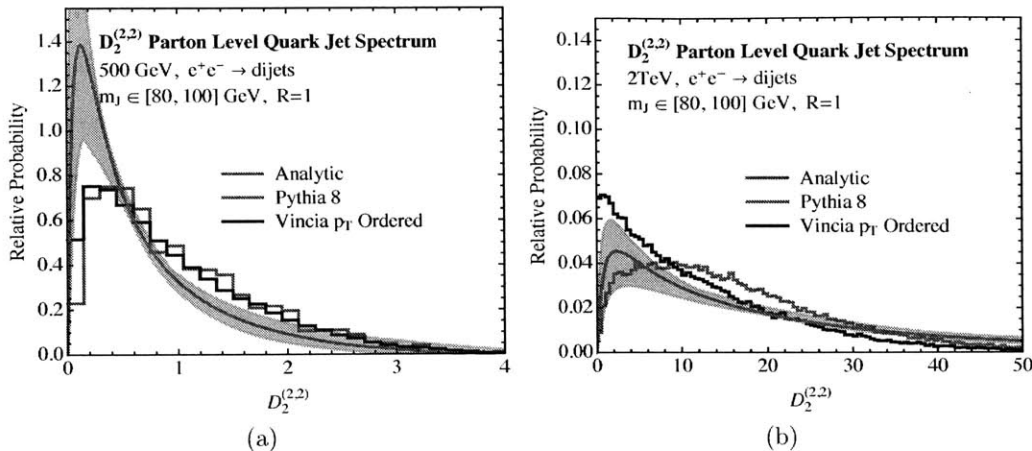
Figure 4-18: Comparison of QCD background D_2 distributions from p_T-ordered VINCIA and PYTHIA to our analytic prediction as a function of the jet energy, E_J. The values $E_J = 500$ GeV, and 2 TeV are shown in Figures a) and b), respectively. A jet radius of $R = 1$ is used for all values of the jet energy.

is valid throughout the range of energies we consider. The case when $e_2^{(2)} \lesssim 1$, but not parametrically so, is outside the scope of this chapter.

In Fig. 4-18 we show distributions for the D_2 observable as obtained from Monte Carlo simulation, and compared with our analytic calculation. As in Sec. 4.5.3, we restrict to p_T-ordered VINCIA and PYTHIA at parton level. The perturbative scale variations for each energy value are less extensively explored and are only meant to provide a rough estimate of the perturbative uncertainty. The evolution of the difference between the VINCIA and PYTHIA generators is again quite fascinating, with the discrepancy between the generators increasing significantly with energy, to the point that at 2 TeV the qualitative shape of the distributions doesn't agree. In particular, the behavior at small D_2 is completely different between the two generators, with VINCIA having a large peak, which is not present in PYTHIA.

As discussed in Sec. 4.5.2, this discrepancy between the generators is evidence that wide angle soft radiation is not accurately modeled by the PYTHIA generator, as compared with VINCIA. The phase space for wide angle soft radiation drastically increases as the energy

176

of the jet is increased, with a fixed jet mass and radius. As evidence that this is indeed the cause of the discrepancy, we have checked that the conclusions of Sec. 4.5.3 remains true at higher energy, as long as the jet radius is taken to scale as $R \sim 2m_J/p_T$, so that it constrains the wide angle soft radiation. For example, for $R = 0.2$ at 2 TeV, we find excellent agreement between the D_2 distributions as generated by PYTHIA and VINCIA.[16] However, as was emphasized in Sec. 4.5.2, since this study is performed at parton level, differences in the cutoff of the perturbative shower between the different generators also contribute to this difference. Because the fact that the disagreement is so large between the generators, and is arising from the modeling of soft radiation, this may be an excellent observable to study soft radiation and color coherence in parton showers.

As a reference, in App. A.9 we show distributions of the $e_2^{(2)}$ observable, measured at both 500 GeV, and 2 TeV for both the VINCIA and PYTHIA Monte Carlos, and at both parton and hadron level. Unlike for the D_2 observable, since $e_2^{(2)}$ is set by a single emission, excellent agreement is observed for the $e_2^{(2)}$ observables between PYTHIA and VINCIA both at parton level, emphasizing that D_2 offers a more differential probe of the perturbative shower than single emission observables.

Our analytic predictions at 2 TeV, as shown in Fig. 4-18, are intermediate between the PYTHIA and VINCIA results. They exhibit a peaked structure at small values of D_2, but not to the extent seen in the VINCIA distribution. We believe that this is largely due to the normalization of the distributions, and the fact that we do not match to fixed order in the tail of the distribution. Since this tail becomes longer at higher energies, a larger disagreement in the peak region is also seen. However, the shape agrees qualitatively with the VINCIA result. On the other hand, at 500 GeV, our analytic prediction has a large peak. This is evidence that because the D_2 spectrum is much more sharply peaked at 500 GeV, higher order resummation may be more important in the peak region. However, the relatively good agreement between analytics and Monte Carlo shows that our factorization theorem is able to accurately capture the energy dependence over a large range of energies.

The results for both the jet radius and jet energy dependence of the D_2 distributions

---

[16]We include this plot in App. A.9, Fig. A-5, for reference.

demonstrate that the extent to which the perturbative parton shower is able to describe QCD jets in the two-prong regime depends strongly on the parameters of the jet. In some regions of the phase space, there is not even a qualitative agreement between different generators. This is important to keep in mind for jet substructure studies based solely on a single Monte Carlo generator, and also emphasizes the importance of analytic calculations for jet substructure, in particular in regions of phase space where there is considerable disagreement between the Monte Carlo generators. It is important to note that hadronization will remove some of the discrepancies in the D_2 distributions between the VINCIA and PYTHIA generators, especially at high energies, where it will smear out the peak at low values of D_2. While this improves qualitatively the behavior of the distributions, discrepancies in the shape still remain. This will be discussed in detail in Sec. 4.5.5, along with its incorporation into our analytic calculation. For comparison to precision analytic calculations and interpreting data, it is vital that Monte Carlo generators provide accurate descriptions of both the perturbative and non-perturbative aspects of QCD jets, and not compensating for perturbative discrepancies by the tuning of non-perturbative parameters. This is especially important for disentangling non-perturbative effects from perturbative effects, the latter of which should in principle be under much better control, and for extracting reliable information about non-perturbative QCD from jet physics.

Throughout the rest of the chapter we will focus on jets with radius $R = 1$ at a center of mass energy of 1 TeV.

4.5.5 Impact of Hadronization

Hadronization plays an important role in a complete description of any jet observable, and a description of non-perturbative effects, preferably from a field-theoretic approach, is required to compare with experimental data. An advantage of the factorization approach taken in this chapter is that it allows for a clean separation of perturbative and non-perturbative physics. Non-perturbative effects enter the factorization theorems presented in Sec. 4.3 through the soft function, which describes the dynamics of soft radiation, both perturbative and non-

perturbative, between the jets. For a large class of additive observables, the treatment of non-perturbative physics in the soft function is well-understood, and can be incorporated using shape functions [147, 148, 193, 194, 195]. Shape functions have support over a region of size Λ_{QCD}, and are convolved with the perturbative soft function. In the tail region of the distribution, where the observable is dominated by perturbative emissions, they reduce to a shift. For a large class of observables, this shift is determined by a universal [196, 197] non-perturbative parameter multiplied by a calculable, observable dependent number [197, 198, 199]. Similar shape functions have also been used to incorporate the effects of pile-up and the underlying event at hadron colliders [200].

The effect of non-perturbative physics on multi-differential cross sections has not been well-studied. For the double differential cross section of two angularities, Ref. [88] considered using uncorrelated shape functions for each angularity individually, but it is expected that a complete description would require a shape function incorporating non-perturbative correlations between observables. For the particular case of the D_2 observable, we will argue that a single parameter shape function can be used to accurately describe the dominant non-perturbative effects, and in particular, that a study of multi-differential shape functions with non-perturbative correlations, is not required.

In Sec. 4.4.3 we performed a study of the fixed order singular structure of the D_2 observable in the presence of a jet mass cut. Importantly, we showed that D_2 only has a singularity at $D_2 = 0$, with its behavior at all other values regulated by the mass cut. Non-perturbative corrections to the D_2 observable will play an important role only when the soft scale becomes non-perturbative, which as just argued, for a perturbative mass cut of the form studied in this chapter, only occurs as $D_2 \to 0$. Recall that the D_2 observable is defined as

$$D_2^{(\alpha,\beta)} = \frac{e_3^{(\alpha)}}{(e_2^{(\beta)})^{3\alpha/\beta}}, \qquad (4.61)$$

which is not additive. However, in the two-prong region of phase space, namely $D_2 \to 0$, the value of $e_2^{(\beta)}$ is set to leading power by the hard splitting, and so D_2 effectively reduces to an additive observable. In this region of phase space the description of non-perturbative

179

effects in terms of a shape function can therefore be rigorously justified from our factorization theorem, and it can be applied directly to the D_2 distribution. For large values of D_2, it is not additive, and the use of a shape function cannot be formally justified. However, in this region, a shape function is not required, as any singular behavior is regulated by a mass cut. We therefore will use a shape function that falls off exponentially at large values of D_2. We believe that this is a self-consistent approach until non-perturbative corrections to multi-differential cross sections are better understood.

In the two-prong region of phase space, we have shown that two distinct factorization theorems, namely the soft subjet and collinear subjets, are required, and in Sec. 4.4.4 we showed how these two descriptions can be merged to provide a complete description of the two-prong region of phase space. Importantly, the two factorization theorems describing the two-prong region of phase space have soft functions with different numbers of Wilson lines. The collinear subjets soft function is a two-eikonal line soft function, while the soft subjet soft function has three eikonal lines. Since the shape function describes the non-perturbative contribution to the soft function, in general we should allow for two distinct shape functions, with independent parameterizations. The zero-bin merging procedure in Sec. 4.4.4 would then be performed on the non-perturbative cross sections, after convolution with the appropriate shape function. However, at the level of perturbative accuracy which we work, and because we will simply be extracting our shape function parameters by comparing to Monte Carlo, the use of distinct parameterizations of different shape functions for both the soft subjet and collinear subjets soft functions would introduce many redundant parameters. To simplify the situation in this initial investigation, we will choose to use the same parametrization of the shape function, and the same non-perturbative parameters for both soft functions. This allows for the non-perturbative corrections to be described by a single parameter, and as we will see provides an excellent description of the Monte Carlo data. Because we use the same shape function for both the soft subjet, and collinear subjets soft functions, it also implies that the shape function can be applied after the zero bin merging procedure, namely, directly at the level of the D_2 distribution.

As a simple parametrization of a shape function for D_2, we follow Ref. [200] and consider

$$F(\epsilon) = \frac{4\epsilon}{\Omega_D^2} e^{-2\epsilon/\Omega_D}, \qquad (4.62)$$

where ϵ is the energy and $\Omega_D \sim \Lambda_{\text{QCD}}$ is a non-perturbative scale. Note that while we will use the same value of Ω_D for the signal and background distributions, it will have very different effects on the two distributions, which will arise naturally from the power counting in the different factorization theorems, as will be shown in this section. The function of Eq. (4.62) satisfies the required properties that it is normalized to 1, has a finite first moment Ω_D, vanishes at $\epsilon = 0$, and falls off exponentially at high energies [194]. More general bases of shape functions are discussed in Ref. [195], although we find that the single parameter shape function of Eq. (4.62) is sufficient to describe the dominant effects of hadronization.

As discussed above, we will use the shape function of Eq. (4.62) for both the collinear subjets and soft subjets factorization theorems, with the same value of Ω_D in both cases. Because we have enforced this simplification to reduce the number of parameters, it is then most interesting to focus on Ω_D for the collinear subjets factorization theorem, which has two eikonal lines. In this case, we can show that we can relate the Ω_D parameter to universal non-perturbative parameters appearing in $e^+e^- \to$ dijet factorization theorems, which have been measured in experiment. Therefore, throughout the rest of this section, we will focus on deriving scaling relations for Ω_D, assuming we are working in the collinear subjets factorization theorem. Again, we wish to emphasize that this is merely a simplification we have made to reduce the number of parameters, and a more general treatment could be performed, but we will see that with only the single Ω_D, with properties derived assuming the collinear subjets factorization theorem, excellent agreement with Monte Carlo is observed.

The effect of non-perturbative physics as modeled by the shape function is very different for background or signal distributions. For background, when D_2 is small, the contribution to $e_3^{(\alpha)}$ from a non-perturbative soft emission is

$$e_3^{(\alpha)}\Big|_{\text{np}} \sim \frac{\epsilon}{E_J} e_2^{(\alpha)}, \qquad (4.63)$$

181

where ϵ is the energy of the non-perturbative emission and E_J is the energy of the jet, as shown in Eq. (4.15). The non-perturbative contribution to D_2 is therefore

$$D_2^{(\alpha,\beta)}\Big|_{\mathrm{np}} = \frac{e_3^{(\alpha)}\big|_{\mathrm{np}}}{(e_2^{(\beta)})^{3\alpha/\beta}} \sim \frac{\epsilon}{E_J}\frac{e_2^{(\alpha)}}{(e_2^{(\beta)})^{3\alpha/\beta}} . \qquad (4.64)$$

In terms of the shape function, the non-perturbative distribution of D_2 for background jets can then be written as a convolution:[17]

$$\frac{d\sigma_{\mathrm{np}}}{dD_2^{(\alpha,\beta)}} = \int_0^\infty d\epsilon\, F(\epsilon)\, \frac{d\sigma_{\mathrm{p}}\left(D_2^{(\alpha,\beta)} - \frac{\epsilon}{E_J}\frac{e_2^{(\alpha)}}{(e_2^{(\beta)})^{3\alpha/\beta}}\right)}{dD_2^{(\alpha,\beta)}} , \qquad (4.65)$$

where σ_{np} and σ_{p} denote the non-perturbative and perturbative cross sections, respectively.

We can estimate the scale at which the global softs of the collinear subjets factorization theorem become non-perturbative from the scaling of the modes given in Eq. (4.14). Rewriting this scaling in terms of the center of mass energy of the e^+e^- collision, Q, and D_2, we find that the global soft scale of the collinear subjets factorization theorem has virtuality

$$\mu_S = 2^3 D_2\, m_Z \left(\frac{m_Z}{Q}\right)^3 , \qquad (4.66)$$

where we have assumed a jet mass, $m_J = m_Z$, as relevant for boosted Z discrimination. Taking $\Lambda_{\mathrm{QCD}} = 500$ MeV, we find that the global soft scale enters the non-perturbative regime at $D_2 \simeq 1$.

Restricting to $\beta = 2$, in the collinear subjets region of the background jet phase space, the non-perturbative distribution of $D_2^{(\alpha,2)}$ is then

$$\frac{d\sigma_{\mathrm{np}}}{dD_2^{(\alpha,2)}} = \int_0^\infty d\epsilon\, F(\epsilon)\, \frac{d\sigma_{\mathrm{p}}\left(D_2^{(\alpha,2)} - 2^{\alpha-2}\frac{\epsilon}{E_J}\frac{E_J^{2\alpha}}{m_J^{2\alpha}}\right)}{dD_2^{(\alpha,2)}} , \qquad (4.67)$$

---

[17]In this initial investigation we do not include a gap in our shape function, which would implement a minimum hadronic energy deposit, as expected physically [194]. Such gapped shape functions, and their associated renormalon [201] ambiguity [202] have been studied for arbitrary angular exponents [203], and could be straightforwardly incorporated in our analysis. However, we observe excellent agreement with our single parameter shape function, which we therefore find to be sufficient for our purposes.

where we have used

$$e_2^{(2)} = \frac{m_J^2}{E_J^2} , \qquad (4.68)$$

and that, in the collinear subjets region of phase space,

$$e_2^{(\alpha)} \simeq 2^{\alpha-2} \left(e_2^{(2)} \right)^{\alpha/2} . \qquad (4.69)$$

Because we consider fixed-energy jets with masses in a narrow window, $e_2^{(2)}$ is just a number and can be removed by appropriate change of variables. Making this change, we then have

$$\frac{d\sigma_{\mathrm{np}}}{dD_2^{(\alpha,2)}} = \int_0^\infty d\epsilon \, F(\epsilon) \, \frac{d\sigma_{\mathrm{p}} \left(D_2^{(\alpha,2)} - \frac{\epsilon}{E_J} \right)}{dD_2^{(\alpha,2)}} , \qquad (4.70)$$

where the non-perturbative parameter in the shape function is effectively modified to

$$\tilde{\Omega}_D = 2^{\alpha-2} \frac{\Omega_D}{\frac{m_J^{2\alpha}}{E_J^{2\alpha}}} . \qquad (4.71)$$

The non-perturbative parameter Ω_D still has implicit dependence on the angular exponent α. Because the global soft modes have the lowest virtuality and can only resolve the back-to-back soft Wilson lines in the n and $\bar{n}$ directions, we can use the results of Refs. [198, 199] to extract the α dependence. By the boost invariance of the soft function[18] along the $n - \bar{n}$ directions and the form of the observable $e_3^{(\alpha)}$ as measured on soft particles, it follows that Ω_D takes the form

$$\Omega_D = \frac{3}{2\alpha - 1} \Sigma , \qquad (4.72)$$

where Σ is a universal non-perturbative matrix element of two soft Wilson lines and all dependence on α has been extracted.[19] We have normalized the matrix element such that

---

[18]This boost invariance holds strictly only for a soft function with no jet algorithm restrictions. However, since we are considering fat jets close to hemispherical, we expect corrections to the boost invariance of the soft function to be small.

[19]In this section we ignore the effects of hadron masses, and their associated power corrections of $\mathcal{O}(m_H/Q)$, where m_H is the mass of a stable hadron in the jet. While these power corrections can also be incorporated through the shape function, in general, they break the universality of the non-perturbative matrix element, Σ [155, 156]. In particular, Eq. (4.72) is no longer in general true, for a Σ that is independent of the angular

the coefficient is unity for $\alpha = 2$. We will shortly discuss the extent to which the values of Ω_D we obtain from comparison with the parton shower agree with the known values of this universal non-perturbative matrix element.

For signal jets, the lowest virtuality mode in the jet are the collinear-soft modes. Unlike the global soft modes of the collinear subjets factorization theorem, which did not resolve the substructure of the jets, allowing us to relate the non-perturbative parameter appearing in the shape function to that appearing in dijet event shapes, the collinear soft modes in the signal factorization theorem resolve the jet substructure. However, since the decaying boson is a color singlet, there are still only two eikonal lines present in the factorization theorem. Boost invariance of the soft function will therefore again allow us to relate the non-perturbative parameter for the signal distribution to that appearing in dijet event shapes. This is similar to the argument used in Ref. [85] to calculate the signal distribution for 2-subjettiness.

A non-perturbative collinear-soft emission contributes to $e_3^{(\alpha)}$ as

$$e_3^{(\alpha)}\Big|_{\mathrm{np}} \sim \frac{\epsilon}{E_J}(e_2^{(\alpha)})^3 \,, \tag{4.73}$$

where now ϵ is the energy of the non-perturbative collinear-soft emission, as shown in Eq. (4.16). The non-perturbative contribution to D_2 for signal jets is therefore

$$D_2^{(\alpha,\beta)}\Big|_{\mathrm{np}} = \frac{e_3^{(\alpha)}\Big|_{\mathrm{np}}}{(e_2^{(\beta)})^{3\alpha/\beta}} \sim \frac{\epsilon}{E_J}\frac{(e_2^{(\alpha)})^3}{(e_2^{(\beta)})^{3\alpha/\beta}} \tag{4.74}$$

$$\simeq 2^{3(\alpha-\beta)}\frac{\epsilon}{E_J}\,,$$

where in the second line we have used the parametric relationship between $e_2^{(\alpha)}$ and $e_2^{(\beta)}$ in the collinear subjets region. Convolving with the shape function, the non-perturbative

exponent α [155, 156]. This depends on the precise definition of the energy correlation functions for massive particles. However, the value of Σ can still be extracted from dijet event shapes in the same universality class as a particular angularity [156]. Furthermore, Ω_D has a scale dependence from renormalization group evolution, $\Omega_D = \Omega_D(\mu)$, although this dependence is logarithmic, and is therefore small compared to our uncertainties. We will discuss briefly the impact of hadron masses and the renormalization group evolution of Ω_D in Sec. 4.6, and in App. A.8.

distribution for signal jets is then

$$\frac{d\sigma_{\text{np}}}{dD_2^{(\alpha,\beta)}} = \int_0^\infty d\epsilon \, F(\epsilon) \, \frac{d\sigma_{\text{p}} \left(D_2^{(\alpha,\beta)} - 2^{3(\alpha-\beta)} \frac{\epsilon}{E_J} \right)}{dD_2^{(\alpha,\beta)}} \, . \tag{4.75}$$

It is important to note how the different scales for the soft radiation in the case of the signal and background jets leads to different behavior of the D_2 distributions after hadronization. In particular, from Eqs. (4.70) and (4.75) one can determine the shift in the first moment of the D_2 distribution caused by hadronization, which we will denote by Δ_D. Restricting to the case $\alpha = \beta = 2$ for simplicity, we find that for the background distribution,

$$\Delta_D = \frac{\Omega_D}{E_J \left(\frac{m_J}{E_J}\right)^4} \, , \tag{4.76}$$

whereas for the signal jets, we have

$$\Delta_D = \frac{\Omega_D}{E_J} \, . \tag{4.77}$$

Since Ω_D should be of the scale 1 GeV, we see that for signal jets, the shift in the first moment due to hadronization is highly suppressed, and behaves differently than a traditional event shape due to the boost factor, while for background jets, since $m_J \ll E_J$, the effect of hadronization is significant. We will see that both of these features, which are consequences of the power counting of the dominant modes, are well reproduced in the Monte Carlo simulations.

Comparisons between the hadron-level distributions of $D_2^{(2,2)}$ from our analytic calculations and the Monte Carlos are presented in Fig. 4-19 for background and Fig. 4-21 for signal jets. For background distributions, we compare our perturbative calculation convolved with the shape function, as defined in Eq. (4.65). Both VINCIA and PYTHIA use the same hadronization model, but HERWIG++ uses a distinct hadronization model, and therefore we allow for a different shape parameter, Ω_D, for the two cases. For the case of PYTHIA and VINCIA, because we find the best agreement in the shape of the perturbative spectrum, with parton level VINCIA with p_T ordering, we choose to extract the value of Ω_D by fitting to the

185
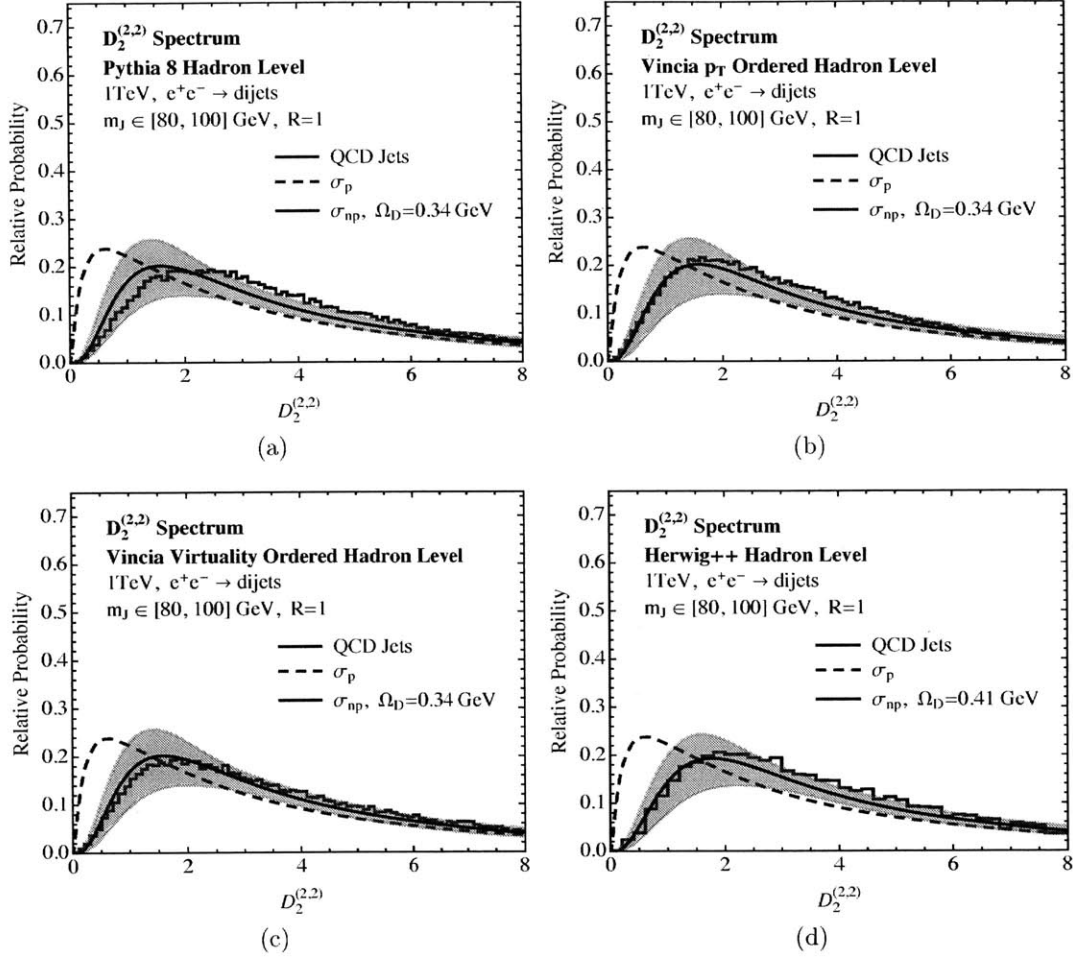
Figure 4-19: A comparison of the $D_2^{(2,2)}$ distributions for background QCD jets from our analytic prediction and the various hadron-level Monte Carlos. σ_p denotes the parton level perturbative prediction for the distribution and $\sigma_{np} = \sigma_p \otimes F_D$ is the perturbative prediction convolved with the non-perturbative shape function. The values of the non-perturbative parameter Ω_D used are also shown.

hadronized distribution for p_T ordered VINCIA. However, we will shortly discuss the level of ambiguity in Ω_D arising from this extraction. For jets with an energy of 500 GeV and mass of 90 GeV, we find that the choice $\Omega_D = 0.34 \pm 0.03$ GeV provides the best agreement of our perturbative calculation with p_T ordered VINCIA, while $\Omega_D = 0.41 \pm 0.03$ GeV provides the best agreement with HERWIG++. The errors assigned here come only from the fitting itself, and are due to the statistical uncertainties of the Monte Carlo distributions due to the finite width of the histogram bins. These errors do not take into account any other uncertainties; for example, whether one should perform the fit to hadron level VINCIA or PYTHIA. This level of agreement between the non-perturbative parameters extracted from PYTHIA and HERWIG++ is comparable to more detailed studies, such as Ref. [156]. A comparison of the distributions of Fig. 4-19 before and after hadronization shows that hadronization has a considerable effect on the background distributions, particularly at small values of D_2, as expected from Eq. (4.76). This effect, which in the Monte Carlos is realized through tuned hadronization models, is well described by the single parameter shape function. Importantly, as discussed above, if different shape parameters were used for the collinear subjets and soft subjets factorization theorems, they would be nearly degenerate in the fit at the level of perturbative accuracy that we work, which is why we have made the simplification of working with a single non-perturbative shape parameter.

We have argued that the non-perturbative parameter Ω_D in the collinear subjets factorization theorem can be related to a universal non-perturbative matrix element of two soft Wilson lines. Such non-perturbative matrix elements appear in the factorization theorems of a large class of e^+e^- event observables, and has therefore been measured from data at LEP.[20] While the value of Ω_D that we have determined for the two parton showers is by no means precise, it is interesting to compare our value with those extracted from precision studies of e^+e^- collider observables which have been performed in the literature. Using the particular case of $\alpha = \beta = 2$, and converting to our normalization, a recent extraction of the

---

[20]An extremely large literature exists on such measurements, and their theoretical interpretation, to which we cannot do justice in this brief section. We refer the reader to, for example, Refs. [147, 148, 204, 205, 2, 206, 1, 207] and references therein.
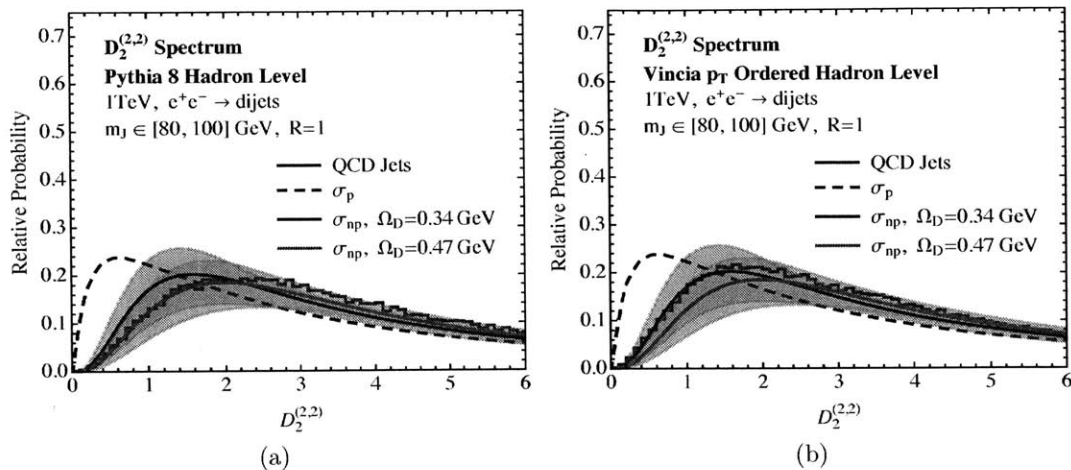
Figure 4-20: A comparison of the $D_2^{(2,2)}$ distributions for background QCD jets from our analytic prediction and PYTHIA and p_T ordered VINCIA Monte Carlos in a) and b). Analytic predictions for different values of the non-perturbative shape parameter Ω_D are shown.

non-perturbative parameter from an N^3LL' analysis of the C-parameter event shape using LEP data, and including power corrections and hadron mass effects [155, 156], gives a value of $\Omega_D = 0.28$ GeV [1, 207]. This agrees well with our values extracted through comparison with Monte Carlo. Going forward, with the goal of increasing both the precision and understanding of jet substructure, the ability to relate the dominant non-perturbative corrections to the D_2 observable to known non-perturbative parameters measured in e^+e^- is a valuable feature, and that further study on the non-perturbative corrections to multi-differential cross sections is of great importance.

Many of the features of the background distributions which were present before hadronization in Fig. 4-14a persist after convolution with the shape function. However, they are greatly reduced, and they become difficult to disentangle from modifications to the non-perturbative shape parameter at the order we work. In particular, from Fig. 4-19, we see that for the choices of Ω_D that we have used, both VINCIA showers agree well with our analytic calculation. On the other hand, there is significant disagreement between our calculation with the chosen values of Ω_D and PYTHIA. The D_2 distribution in PYTHIA is systematically pushed to higher values as compared with our calculation.

To try and asses the extent to which this can be accommodated for by adjusting the value of Ω_D, in Fig. 4-20 we show plots of both PYTHIA and VINCIA with p_T ordering compared with our analytic results for two different values of the shape parameter. The values $\Omega_D = 0.34$ GeV and $\Omega_D = 0.47$ GeV were chosen to give best agreement with the VINCIA and PYTHIA distributions, respectively. This figure makes clear that the disagreement between the D_2 distributions as generated by the two Monte Carlo generators can largely be remedied by using different values of the non-perturbative parameter. We note also that the effect of changing the non-perturbative parameter is of course similar to that of changing the perturbative cutoff of the shower, as was discussed in Sec. 4.5.2, making it difficult to disentangle these two effects. However, we have argued that there are also legitimate differences in the modeling of soft radiation between the generators. While these can be partially compensated for by non-perturbative physics, with higher perturbative accuracy, one could begin to disentangle these effects.

This plot also gives a feel for the extent to which Ω_D can be varied before significant disagreement is seen between the analytic calculation and a given Monte Carlo distribution. Performing the perturbative calculation to higher accuracy would help to resolve some of these ambiguities in the value of the shape parameter, by reducing the perturbative uncertainty on the shape of the distribution, as well as its normalization. Throughout the rest of this chapter, when comparing our analytic predictions with VINCIA or PYTHIA, we will use the value $\Omega_D = 0.34$ GeV as obtained from our fit to hadron level p_T-ordered VINCIA. From comparison to our analytic calculations, we believe that VINCIA best describes the perturbative distribution and implements a clean separation between perturbative and non-perturbative physics, as in our analytic calculation. However, one should keep in mind the level of sensitivity to this parameter. In particular, for the application of boosted Z discrimination, we will see that the discrimination power of the observable will depend sensitively on the shape of the D_2 distribution below the peak, and will therefore exhibit great sensitivity to the value of the shape parameter.

For the signal distributions, shown in Fig. 4-21, we use the same choice of non-perturbative

189
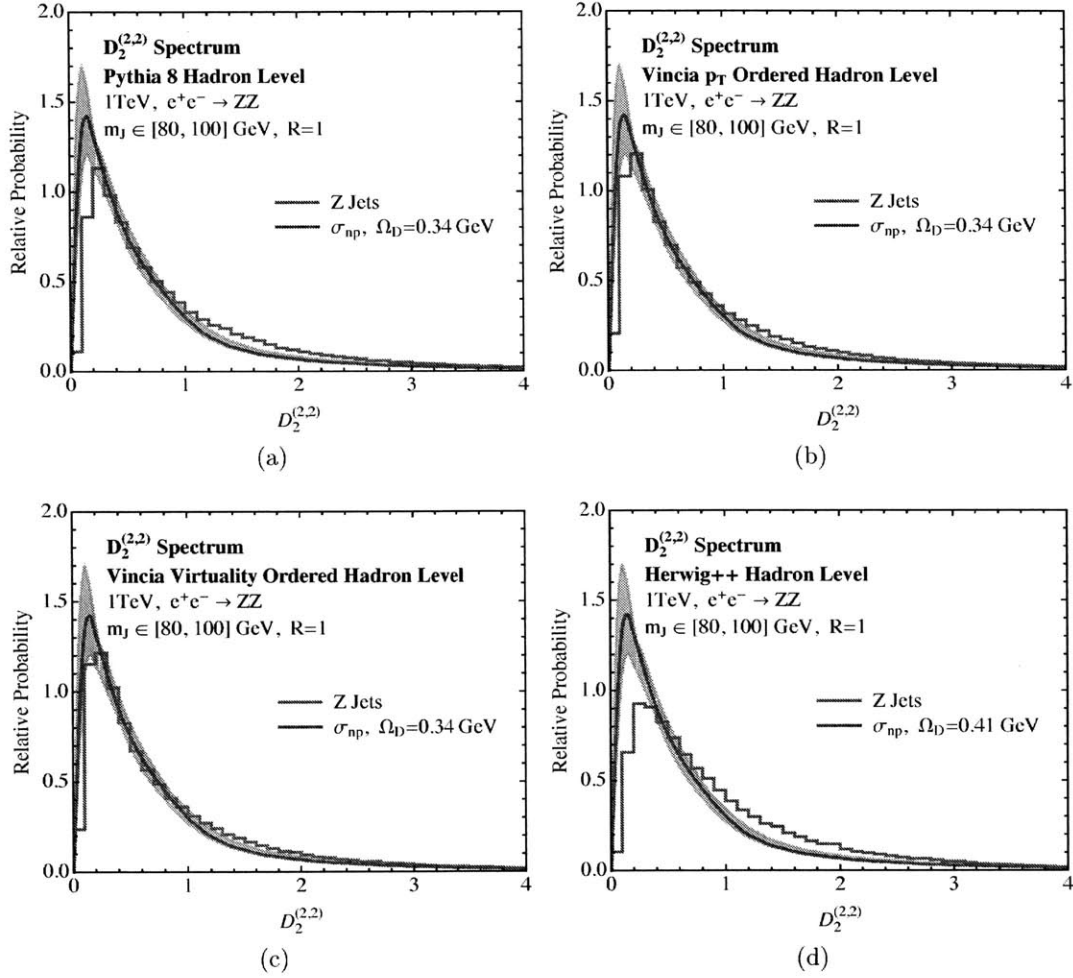
Figure 4-21: A comparison of the $D_2^{(2,2)}$ distributions for signal boosted Z jets from our analytic prediction and the various hadron-level Monte Carlos. σ_p denotes the parton level perturbative prediction for the distribution and $\sigma_{np} = \sigma_p \otimes F_D$ is the perturbative prediction convolved with the non-perturbative shape function, although for the signal this has a negligible effect. The values of the non-perturbative parameter Ω_D used are also shown.

parameters as for the background distributions. From Eq. (4.77), we have seen that for the jets with $E_J = 500$ GeV, the non-perturbative shift is expected to be of the order $1/500$, and is therefore completely negligible to the level of accuracy that we work, and the equality of the non-perturbative parameters between the signal and background distributions is not tested. For the signal distributions, we see excellent agreement between the theory prediction and all the Monte Carlo generators. Due to the sharp peak in the distribution, we expect higher order resummation is necessary to provide a more accurate description right in the peak region, where the perturbative uncertainty in our calculation becomes large. Due to the fact that the distributions are normalized, this uncertainty also manifests itself in the tail of the distribution. It is known how to calculate the signal distribution to higher accuracy [85], and so we do not consider this issue further here. The effect of the shape function on our analytic results are consistent with all of the Monte Carlos, whose signal D_2 distribution is changed only slightly (i.e., only in the lowest bins) after hadronization.

We conclude this section by emphasizing how the choice of variable can greatly facilitate comparisons with Monte Carlos. An important feature of the D_2 observable is that it cleanly separates phase space regions dominated by different physics. In particular, it separates the region of phase space where a subjet is formed from that where no subjet is formed, as well as separating the regions of phase space where hadronization is important from those where it plays a minor role. This enables these effects to be cleanly disentangled, and provides a sensitive probe of their modeling. We therefore believe that the observable D_2 could play an important role in the tuning of Monte Carlo generators for jet substructure studies, and could be used to complement some of the observables proposed and studied in Refs. [177, 178].[21] Furthermore, the observable D_3 [8], which is sensitive to three-prong substructure within a jet also provides a clean separation of two- and three-prong regions, and could be used to provide an even more detailed understanding of jet substructure and the perturbative shower

---

[21]Note that Refs. [177, 178] used the observable C_2, also formed from the energy correlation functions, which was proposed in Ref. [30]. Unlike D_2, C_2 does not cleanly separate the two-prong region of phase space from the one-prong region of phase space. A detailed discussion of this point can be found in Ref. [5]. The clean separation of the one- and two-prong regions of phase space is the essential feature of the D_2 observable, which allows for its precise theoretical calculation and its sensitivity to the shower implementation.

evolution.

4.5.6 Analytic Boosted Z Discrimination with D_2

In this section, we use our analytic calculation, combined with the non-perturbative shape functions of Sec. 4.5.5, to make complete predictions for the discrimination power of $D_2^{(2,2)}$ for hadronically-decaying boosted Z bosons versus QCD quark jets at an e^+e^- collider. We present comparisons of our calculation to the results of fully hadronized PYTHIA, VINCIA, and HERWIG Monte Carlos. Here, we also present Monte Carlo results from scanning over a range of values for the angular exponent α that is consistent with our factorization theorem. Analytic results for boosted boson discrimination were also presented recently in Ref. [133] for groomed mass taggers, as well as an analytic study of the optimal parameters.

In Figs. 4-22 and 4-23 we overlay the distributions for $D_2^{(2,2)}$ as measured on signal and background for each Monte Carlo sample, and compare with our analytical calculations including the non-perturbative shape function contributions. Fig. 4-22 shows the complete D_2 distributions, including the long tail of the background distribution, while Fig. 4-23 shows a zoomed in version, focusing on small values of D_2, as is most relevant for signal versus background discrimination. A representative cut on the D_2 distribution, as could be used to select a relatively pure sample of boosted Z bosons, is also indicated. In general, the agreement between the Monte Carlos, for both signal and background distributions, and our calculation is impressive. This holds true both for the overall shape of the distributions, including the long tail of the background distribution, and for the detailed shape at small values of D_2. It is also important to note that the perturbative uncertainties remain under control, even in the small D_2 region, as seen in Fig. 4-23. The uncertainty bands do not incorporate variations in the non-perturbative parameter Ω_D. There are however, some small deviations between the analytic predictions and the Monte Carlo distributions. The background distribution in PYTHIA is pushed to slightly higher values than our calculation. This implies that the signal versus background discrimination power as predicted with PYTHIA will be overestimated. The most conservative prediction for the signal versus background
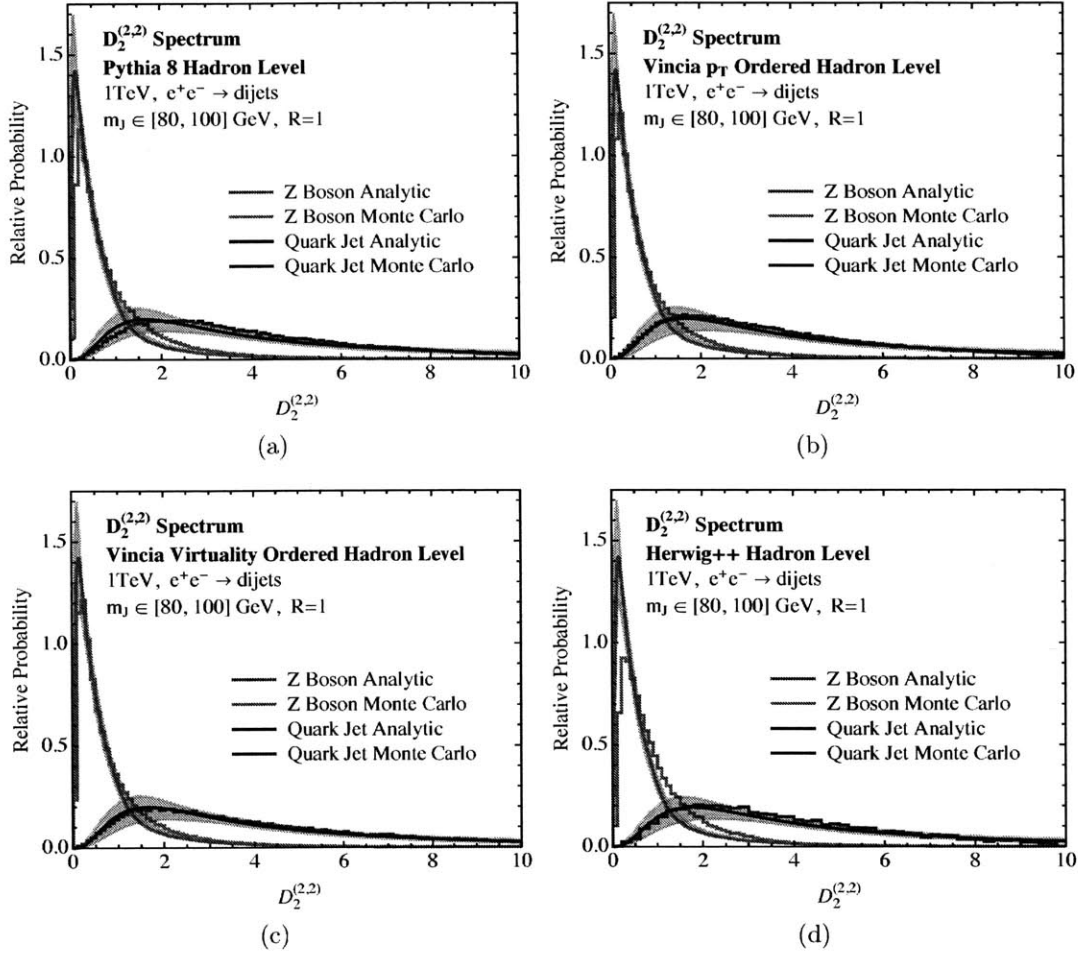
Figure 4-22: A comparison of signal and background $D_2^{(2,2)}$ distributions for the four different Monte Carlo generators and our analytic calculation, including hadronization. Here we show the complete distributions, including the long tail for the background distribution. Although we extend the factorization theorem beyond its naive region of applicability into the tail, excellent agreement with Monte Carlo is found.

193
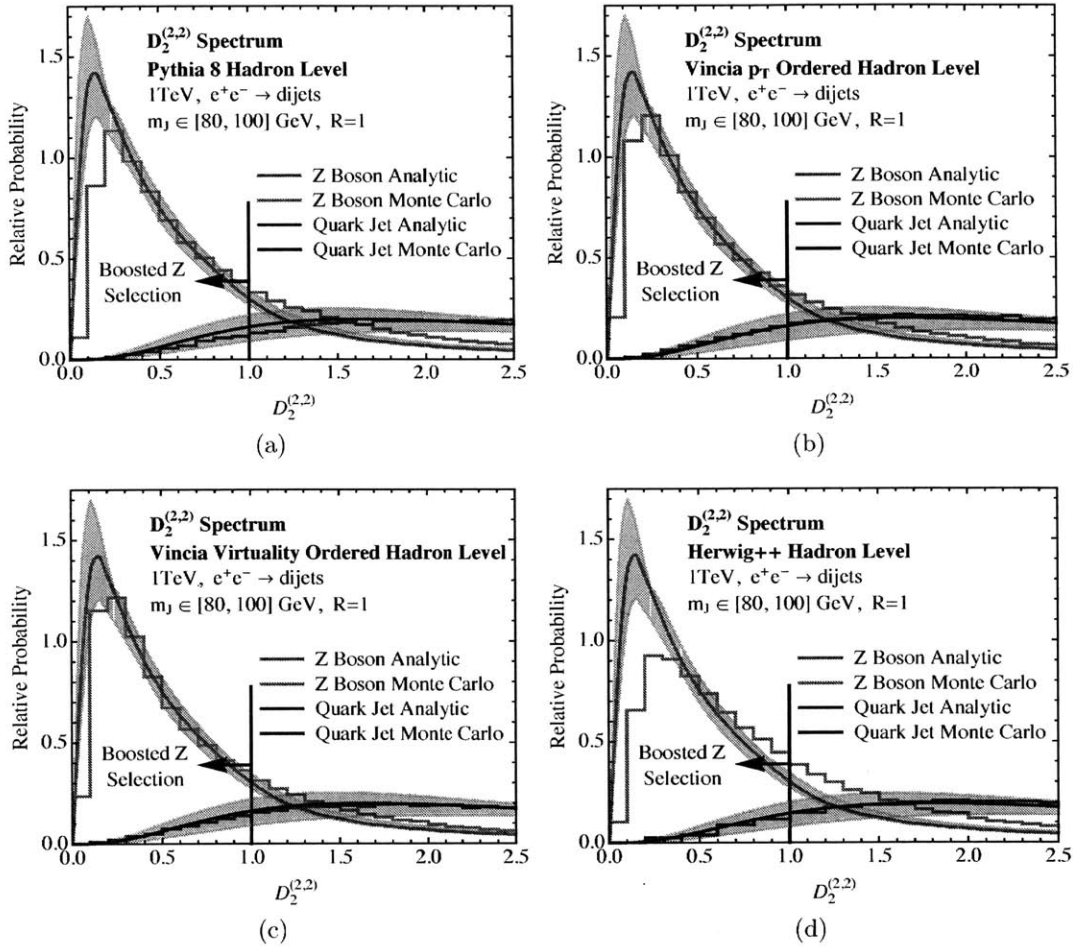
Figure 4-23: A comparison of signal and background $D_2^{(2,2)}$ distributions for the four different Monte Carlo generators and our analytic calculation, including hadronization. Here we show a zoomed in view of the distributions at small D_2, along with a representative cut that could be used to select a relatively pure sample of boosted Z bosons. Relevant cuts for boosted Z discrimination are to the left of the perturbative peak for the background distributions.

194

discrimination power is from HERWIG, whose background distribution is nearly identical to our calculation. That PYTHIA tends to be optimisitic and HERWIG tends to be pessimistic with respect to discrimination power has been observed in several other jet substructure analyses [30, 5, 8, 68].

An important feature of the D_2 distributions, made clear by Fig. 4-23, is that in the region of interest relevant for boosted Z discrimination, the background distribution is deep in the non-perturbative regime. Therefore, although the perturbative uncertainties are small, the effect of the shape function, and variations of the non-perturbative parameter Ω_D, is large. Estimates of the uncertainties due to the form of the shape function, or the use of more complicated functional forms, along the lines of Ref. [195] are well beyond the scope of this chapter. An advantage of our factorization approach is that we are able to achieve a clean separation of perturbative and non-perturbative effects, and demonstrate relations between the non-perturbative matrix elements appearing in our factorization theorems and non-perturbative matrix elements which have been measured with other event shapes, by using their field theoretic definitions. This separation is essential for understanding discrimination performance in the non-perturbative region, which we see is required for jet substructure studies related to boosted boson discrimination. Importantly, though, D_2 seems to take advantage of the different hadronization corrections to signal and background jets, and the overlap of the signal and background regions of D_2 decreases significantly in going from parton-level to fully hadronized jets.

In Fig. 4-24, we have used these raw distributions to produce signal versus background efficiency curves (ROC curves) by making a sliding cut in D_2. The ROC curve from each Monte Carlo sample as well as our analytic prediction from our calculated signal and background distributions are shown in both logarithmic plot and linear plot in Figs. 4-24a and 4-24b, respectively. The band around our analytic prediction should be taken as representative of the signal versus background efficiency range from varying the perturbative scales. [22]

---

[22]Note that ROC curves only make sense for normalized distributions, and therefore the envelopes from scale variation cannot be used. Instead, ROC curves are generated from normalized signal and background distributions made with a variety of scale choices, with scales varied separately in the signal and background distributions. We then take the envelope of these ROC curves to generate the uncertainty bands for the
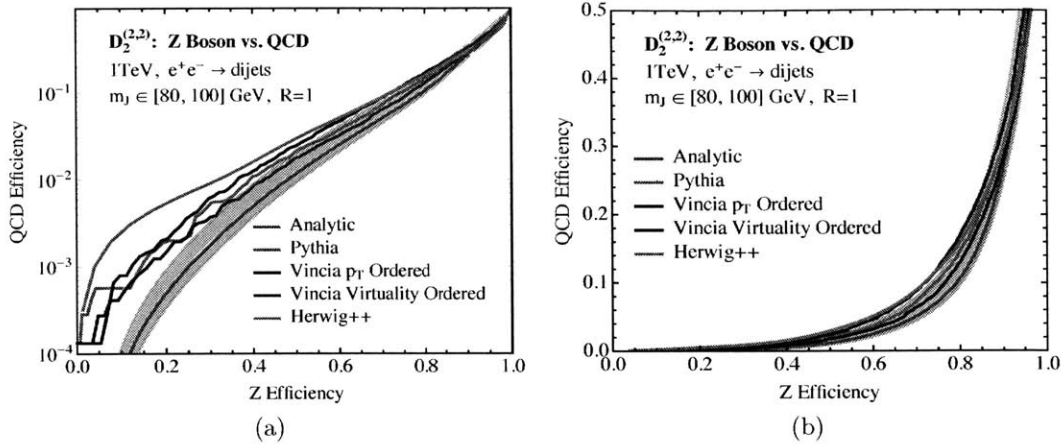
Figure 4-24: Signal vs. background efficiency curves for $D_2^{(2,2)}$ for the Monte Carlo samples as compared to our analytic prediction on a a) logarithmic scale plot and b) linear scale plot. The band of the analytic prediction is representative of the perturbative scale uncertainty.

For the analytic predictions, we use $\Omega_D = 0.34$, as obtained from our fit to the p_T ordered VINCIA shower. Consistent with the distributions in Fig. 4-22, the Monte Carlos are in qualitative agreement with our analytic prediction for the ROC curve. In general, our analytic prediction seems to give an optimistic prediction for the discrimination power, however, this is driven by the fact that our resummed prediction for the signal distribution is more peaked. It would be interesting to perform the NNLL resummation for the signal, which should significantly reduce the uncertainty in the signal calculation, particularly in the peak region, where the perturbative uncertainties in our present calculation are quite large. Because of the fact that the distributions are normalized, an improved behavior in the peak of the distribution could also improve the agreement in the tail of the signal distribution, which is currently systematically low, due to the fact that the peak is systematically high. This could enable a conclusive understanding as to the discrepancy between the different Monte Carlo generators for both signal and background distributions. In particular, our analytic calculations suggest that the HERWIG++ generator provides pessimistic predictions for the discrimination power of the D_2 observable due to the underestimation of the peak height

---

ROC curves.

196
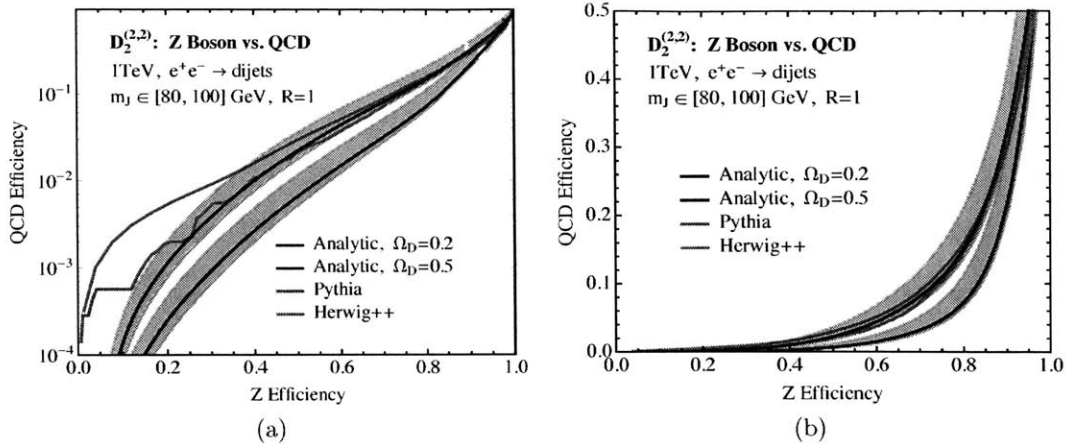
Figure 4-25: Signal vs. background efficiency curves for $D_2^{(2,2)}$ for the Monte Carlo samples as compared to our analytic prediction for two different values of the non-perturbative shape parameter, chosen by varying our central value by ± 0.15 GeV. Results are shown on a logarithmic scale in a) and a linear scale in b). Perturbative scale uncertainties are also shown.

for the signal distribution, and it would be interesting to understand this further. Due to the importance of analytically understanding the discrimination power of jet substructure observables, such a calculation is well motivated. For the case of $\alpha = \beta = 2$, the required perturbative components could be obtained following relations to e^+e^- event shapes as were used in Ref. [85].

One feature made clear by the linear ROC curve in Fig. 4-24b is the increase in perturbative uncertainty with increasing Z efficiency. As emphasized earlier, this is due to the fact that for the region of interest for Z discrimination, one is probing values of D_2 which are below the peak of the background distribution, and therefore in the non-perturbative regime. As the Z efficiency is increased, one enters the peak region of the background distribution, where the perturbative uncertainty is largest, causing a corresponding increase in the uncertainty band for the ROC curve. However, we do not include uncertainties due to the non-perturbative parameter Ω_D or from the shape function, in Fig. 4-24b, which are the dominant sources of uncertainty in this region.

To demonstrate that is indeed the case, in Fig. 4-25 we show ROC curves in both linear
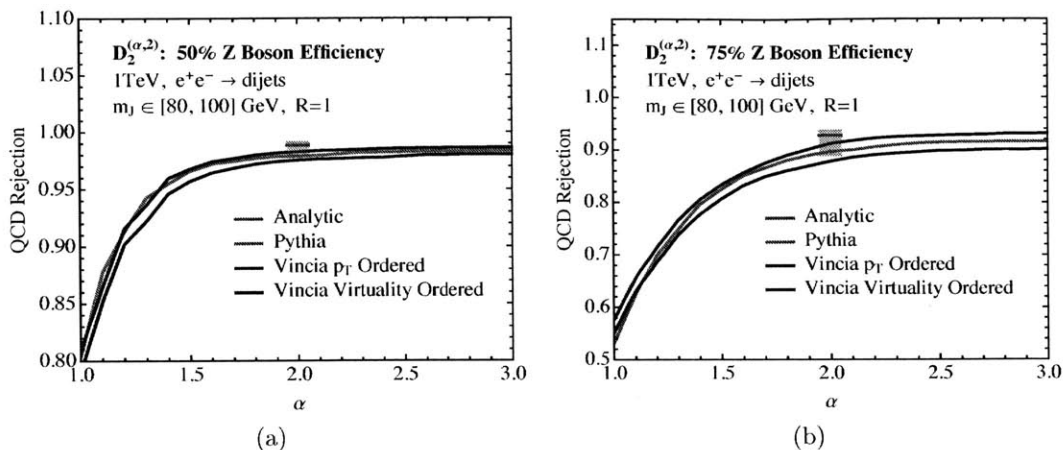
Figure 4-26: Background rejection rate at fixed a) 50% and b) 75% signal efficiency as a function of the angular exponent of the 3-point energy correlation function in D_2, and a comparison to our analytic prediction for $\alpha = 2$.

and log scales for two different values of the non-perturbative shape parameter. The values of Ω_D where chosen by varying our central value of $\Omega_D = 0.34$ GeV by ± 0.15 GeV (and rounding to nice numbers). We have also shown the distributions from the HERWIG++ and PYTHIA generators as representative of the ROC curves generated by the Monte Carlo generators. This figure makes clear that in the region of efficiencies of interest for boosted Z tagging, one is extremely sensitive to the D_2 distribution in the deeply non-perturbative region, and this uncertainty swamps the perturbative uncertainty. To be able to improve the accuracy in this region will require detailed comparisons with Monte Carlo, data, and analytic calculations, to allow for a clean separation of the non-perturbative parameter from perturbative modifications to the shape of the distribution.

To further understand the discrimination power of the D_2 observable, in Figs. 4-26a and 4-26b we show the background rejection rate at 50% and 75% signal efficiency as a function of α, the angular exponent of the 3-point energy correlation function in D_2. Below about $\alpha = 4/3$, all rejection rates dramatically decrease as α decreases, while above about $\alpha = 4/3$, the QCD rejection rate in all Monte Carlo samples is impressively flat. This is consistent with our power counting analysis of the $e_2^{(2)}, e_3^{(\alpha)}$ phase space plane in Sec. 4.2.2 and is a

powerful verification that the Monte Carlos respect the parametric dynamics of QCD.

Although our factorization theorem is valid in the region $\alpha \gtrsim 2$, for $\beta = 2$, in Figs. 4-26a and 4-26b we have only shown the analytic prediction for the value $\alpha = 2$, where we find that it agrees well with the Monte Carlo results, as expected from the agreement of the distributions and ROC curves. For $\alpha > 2$, while our prediction for the background distribution remains accurate (indeed our power counting becomes more valid in this region), the signal distribution becomes extremely sharply peaked, which is difficult to describe, and sensitive to normalization. Due to the fact that this region is also of less phenomenological interest, both because the large angular exponent makes the observable sensitive to pile up contamination, and because both power counting and Monte Carlo analyses indicate that optimal performance is achieved for $\alpha = 2$, we have decided not to focus on this region. It would be potentially interesting to see if higher order resummation would be sufficient to describe the sharply peaked signal distribution in this region, as well as to test the universality of the non-perturbative power corrections.

One further interesting feature of Figs. 4-26a and 4-26b is the correspondence between the perturbative scale variations, and the spread in the curves from the different Monte Carlo generators, which agree well at both 50% and 75%. For the case of p_T-ordered VINCIA as compared with virtuality ordered VINCIA, this correspondence is precise, as the difference between the Monte Carlos can be viewed as a scale variation, and identical hadronization models are used.

4.5.7 Discrimination in the Two-Prong Regime

Throughout this chapter, we have emphasized that the discrimination of boosted hadronically decaying Z bosons (or W or H bosons) from massive QCD jets is effectively a problem of discriminating one- from two-prong jets. We have demonstrated that the observable D_2 is powerful for this goal. However, in the formulation of our factorization theorem for calculating the distribution of D_2, we needed to perform additional 2-point energy correlation function measurements on the jet to separate contributions from soft subjet and collinear sub-
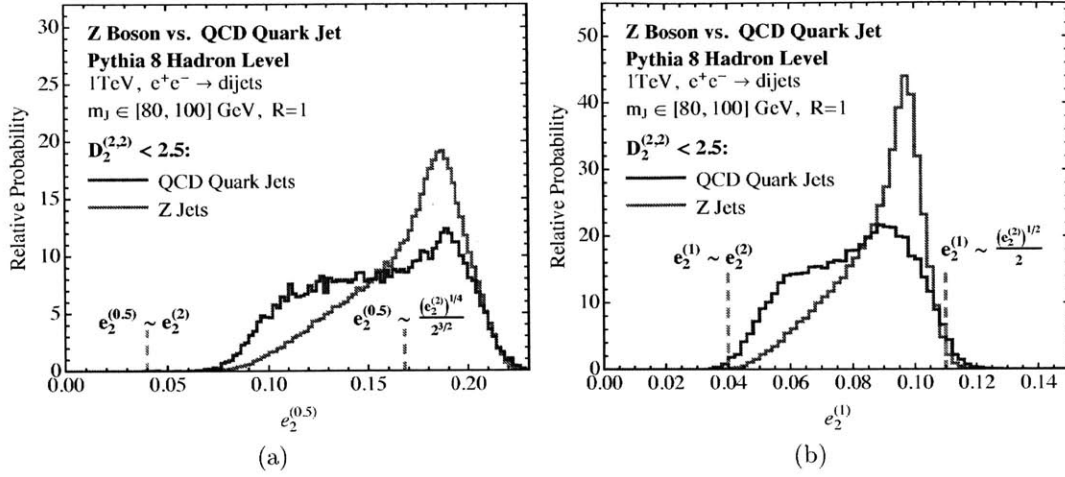
Figure 4-27: Distributions for $e_2^{(0.5)}$ (left) and $e_2^{(1.0)}$ (right) from the signal and background PYTHIA Monte Carlo samples. In addition to the mass cut $m_J \in [80, 100]$ GeV, these jets are also required to have $D_2^{(2,2)} < 2.5$ to guarantee that these jets are dominated by two-prong structure. The parametric boundaries of $e_2^{(\beta)}$ from Eq. (4.78) are shown with the green dashed lines.

jets contributions to background. While indeed the signal jets are dominantly two-pronged, we further know that those prongs are dominantly collinear, and do not have parametrically different energies. Therefore, we are able to further discriminate signal from background jets in the two-prong region of phase space by exploiting additional measurements that can isolate the soft subjet and collinear subjet configurations. A detailed analysis of this is beyond the scope of this chapter, but here, we will demonstrate in Monte Carlo that such a procedure is viable.

To investigate this, we measure the observable $D_2^{(2,2)}$ on jets on which a tight mass window cut has been applied. Other angular exponents for D_2 can be used also, but here we only measure D_2 to define two-prong jets. We restrict to the two-prong region of phase space by requiring that $D_2^{(2,2)} < 2.5$. Then, on the jets that pass these cuts, we measure two, 2-point energy correlation functions, $e_2^{(2)}$ and $e_2^{(\beta)}$, where $\beta < 2$. As discussed in Sec. 4.3, the measurement of the two 2-point energy correlation functions provides an IRC safe definition of the subjets' energy fractions and splitting angle. Because we make a tight mass cut on

the jets, $e_2^{(2)}$ is essentially fixed, and only $e_2^{(\beta)}$ is undefined. We will study the distribution of $e_2^{(\beta)}$ for both signal and background jets in this region of phase space.

For a fixed value of $e_2^{(2)}$ and $\beta < 2$, $e_2^{(\beta)}$ is parametrically bounded as

$$e_2^{(2)} \lesssim e_2^{(\beta)} \lesssim 2^{\beta-2}(e_2^{(2)})^{\beta/2} \,. \tag{4.78}$$

In the two-prong region, the lower bound is set by the soft subjet while the upper bound is set by collinear subjets. Therefore, $e_2^{(\beta)}$ for signal jets will peak near $2^{\beta-2}(e_2^{(2)})^{\beta/2}$, while background QCD jets will fill out the full range. We illustrate this in Fig. 4-27 on the hadronized PYTHIA sample with the appropriate cuts applied. We show plots of the distributions of $e_2^{(0.5)}$ and $e_2^{(1)}$ on both signal and background jets and have added dotted lines to denote the parametric upper and lower boundaries. As expected, signal peaks near the upper boundary and background fills out the entire allowed region and so this additional information could be used for discrimination. For the very small values of $\beta = 0.5$, an $\mathcal{O}(1)$ drift is observed with respect to the parametric boundaries, while for $\beta = 1$, the parametric boundaries are extremely well respected.

This demonstrates a simple example of an observable which goes beyond the simple one vs. two prong picture of jet substructure, asking more differential questions about the subjets themselves. In particular, it could be used both to further improve the discrimination power of boosted boson discriminants, and to study in detail the QCD properties of subjets.

4.6   Looking Back at LEP

In this section, we consider the D_2 distribution for QCD jets in e^+e^- collisions at the Z pole at LEP, for which a large amount of data exists. While the use of D_2 for boosted boson discrimination is not possible, nor relevant, at LEP, this will emphasize the sensitivity of D_2 as a probe of two-prong structure in jets. We will suggest the importance of using variables sensitive to two emissions off of a primary quark in tuning Monte Carlo generators to LEP data.

Our definition of the energy correlation functions in Eq. (4.1) makes implicit assumptions about the treatment of hadron masses, which we have ignored to this point. The definition given there is an E-scheme treatment of hadron masses [155, 156], but we could equally well define p-scheme energy correlation functions as:

$$e_2^{(\beta)} = \frac{1}{E_J^2} \sum_{i<j\in J} |\vec{p}_i| \, |\vec{p}_j| \left[2(1 - \cos\theta_{ij})\right]^{\beta/2} , \qquad (4.79)$$

$$e_3^{(\beta)} = \frac{1}{E_J^3} \sum_{i<j<k\in J} |\vec{p}_i| \, |\vec{p}_j| \, |\vec{p}_k| \left[2(1 - \cos\theta_{ij})2(1 - \cos\theta_{jk})2(1 - \cos\theta_{ik})\right]^{\beta/2} ,$$

where $\vec{p}_i$ denotes the three-momenta of particle i. For massless particles, this definition is identical to that of Eq. (4.1), and so our perturbative analytics would be unchanged by using this definition or the definition of Eq. (4.1).[23] The definitions of Eq. (4.1) and Eq. (4.79) differ for massive particles. In particular, the energy correlation functions as defined in Eq. (4.79) have the advantage that they vanish for low momentum or collimated particles regardless of whether these particles are massless or massive, which is not true of the definition in Eq. (4.1). Because of this, we expect that the energy correlation functions as defined in Eq. (4.79) are less sensitive to hadron mass effects and that kinematic restrictions on the energy correlation functions remain the same before and after hadronization, so that the phase space studied in Sec. 4.2.2 assuming massless particles is not significantly modified.

At LEP energies, hadronization will also have a larger effect on the D_2 spectrum than at 1 TeV. However, a particularly important aspect of our all orders factorization theorem is that it isolates perturbative and non-perturbative physics contributions. In this section we will again implement non-perturbative effects into our analytic calculation using the shape function defined in Eq. (4.62). There are two effects which determine how the shape function depends on the jet mass, m_J, and the center of mass energy, Q. First, for a fixed valued of Ω_D, the shift in the first moment of the D_2 distribution was given in Eq. (4.76), which we

---

[23]As will be discussed shortly, the differences in our analytic calculation due to hadron masses will arise through non-perturbative effects, namely the shape function.

recall here for convenience, by

$$\Delta_D = \frac{\Omega_D}{E_J \left(\frac{m_J}{E_J}\right)^4}.$$
(4.80)

This has dependence on both m_J and Q (through E_J), and for the jets we consider at LEP, this is a considerably larger shift than for the 1 TeV jets studied in Sec. 4.5.5. This scaling is a non-trivial prediction of our factorization framework, and we will see that it is well respected when we perform a comparison of our analytic results with Monte Carlo. Furthermore, the parameter Ω_D has a logarithmic dependence on a renormalization scale, $\Omega_D = \Omega_D(\mu)$, through renormalization group evolution [156], which is briefly reviewed in App. A.8. However, this effect is small compared with the linear change in the first moment with E_J for a fixed m_J/E_J. A numerical estimate for the effect of the running of $\Omega_D(\mu)$ is given in App. A.8. At the level of accuracy to which we work in this chapter, we cannot probe this logarithmic running, although we will see that our results are consistent with it.

The definition of the energy correlation functions given in Eq. (4.79) also has an effect on the universality of the non-perturbative parameter Ω_D, when hadron mass effects are included. Power corrections due to hadron mass effects are of order $\mathcal{O}(m_H/Q)$, where m_H is a light hadron mass, and are therefore of the same order as the leading $\mathcal{O}(\Lambda_{\mathrm{QCD}}/Q)$ power corrections. In the p-scheme definition of the energy correlation functions which we have chosen in Eq. (4.79), it is no longer possible to extract the dependence on the angular exponent alpha from Ω_D, as was done in Eq. (4.72). However, to the accuracy to which we work, we expect this to be a negligible effect, and furthermore, the case $\alpha = 2$ is of most phenomenological interest, and is the case we have focused on exclusively in this chapter. Furthermore, even in the presence of hadron mass effects, it is still possible to extract the parameter Ω_D from dijet event shapes in the same universality class [156]. This exhibits the benefits of the factorization approach both for separating perturbative and non-perturbative effects, and for relating non-perturbative parameters to maintain predictivity.

One further distinction between the case of boosted Z discrimination and the measurement of QCD jet shapes at the Z pole is that while a tight mass cut is natural for boosted Z discrimination, it is not natural in jet shape analyses. However, our shape function analysis,
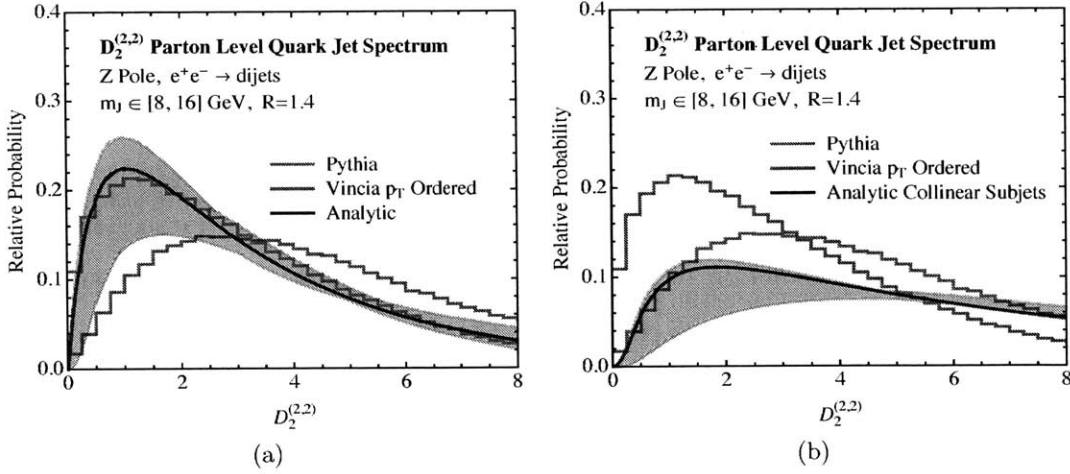
Figure 4-28: A comparison of the D_2 spectrum as measured on quark initiated jets at the Z pole from the PYTHIA and p_T-ordered VINCIA Monte Carlo generators to our analytic predictions. a) Comparison of our complete analytic calculation including both the soft subjet and collinear subjets region of phase space with the predictions of the Monte Carlo generators. b) Comparison of our analytic calculation including only the collinear subjets region of phase space compared with the predictions of the Monte Carlo generators.

as derived in Sec. 4.5.5, is valid at a fixed jet mass (or correspondingly fixed value of $e_2^{(\beta)}$). This is clear from both Eq. (4.67) and from the equation for the shift in the first moment in Eq. (4.76). However, we emphasize that the non-perturbative parameter Ω_D is unique, and the scaling of the non-perturbative shift with the jet mass is fully determined. To achieve an analytic prediction for the non-perturbative D_2 spectrum inclusive over the jet mass mass, one must calculate the perturbative D_2 spectra differentially in the jet mass, convolve with a shape function for each value of the jet mass, and then integrate over the jet mass. While this is in principle straightforward, it is computationally intensive, and is beyond the scope of this chapter. Instead, we will enforce a jet mass cut of $8 < m_J < 16$ GeV. This mass cut was chosen because it is near to the Sudakov peak of the jet mass distribution for this jet energy and the m_J in this range are set by low scale, but still perturbative, emissions.

Similar to what we did in our numerical analysis at 1 TeV, we begin in Fig. 4-28 by comparing our analytic prediction for the D_2 spectrum with the distributions from parton level Monte Carlo. In Fig. 4-28a, we show a comparison of our complete analytic calcula-

204

tion, including perturbative scale variations, along with Monte Carlo predictions from both PYTHIA and p_T-ordered VINCIA, which we take as representative of the different Monte Carlo generators. We use a jet radius of $R = 1.4$ to approximate hemisphere jets. We find good agreement with the prediction with the VINCIA Monte Carlo, and much worse agreement in the shape of the distribution with the PYTHIA generator. We believe that this is again due to the effects of wide angle soft radiation, which are not well modeled by the implementation of the perturbative shower in PYTHIA, but which are well modeled by the antenna dipole shower of the VINCIA Monte Carlo generator, and are explicitly included in our analytic calculation through the soft subjet factorization theorem.[24] To demonstrate that this is indeed the case, in Fig. 4-28b, we show a comparison of our analytic prediction, including only the collinear subjets region of phase space, with both Monte Carlo generators. Excellent agreement is seen with the PYTHIA generator, particularly at small D_2, while the agreement between the prediction including only the collinear subjets region of phase space and the VINCIA prediction are completely different. This emphasizes the large effect played by the soft subjet at LEP energies.

In Fig. 4-29b we show our analytic prediction for the non-perturbative spectra using the shape function. An alternate view of the perturbative spectrum is shown in Fig. 4-29a for reference, and to show the overall shape of the perturbative distribution. We have used a valued of $\Omega_D = 0.50$ GeV, which was obtained by fitting to the VINCIA Monte Carlo, since we obtained the best agreement with the shape of the VINCIA Monte Carlo at parton level with our perturbative spectrum. There is considerable uncertainty on this value, probably of the order ± 0.3 GeV due to the wide mass window, which is probably slightly large for the naïve application of our shape function. Furthermore, as demonstrated in Sec. 4.5.5, there is some ambiguity in the value of Ω_D, depending on whether it is extracted from hadron level PYTHIA or VINCIA, which is of this same order. However, this value is consistent with $\Omega_D = 0.34$ GeV as extracted from our analysis at 1 TeV. Although it is expected that the logarithmic running of the Ω_D parameter will decrease its value slightly, this effect is expected to be

---

[24]Again it is important to emphasize, that, as was discussed in Sec. 4.5.2, there are also large uncertainties due to the treatment of the shower cutoff.
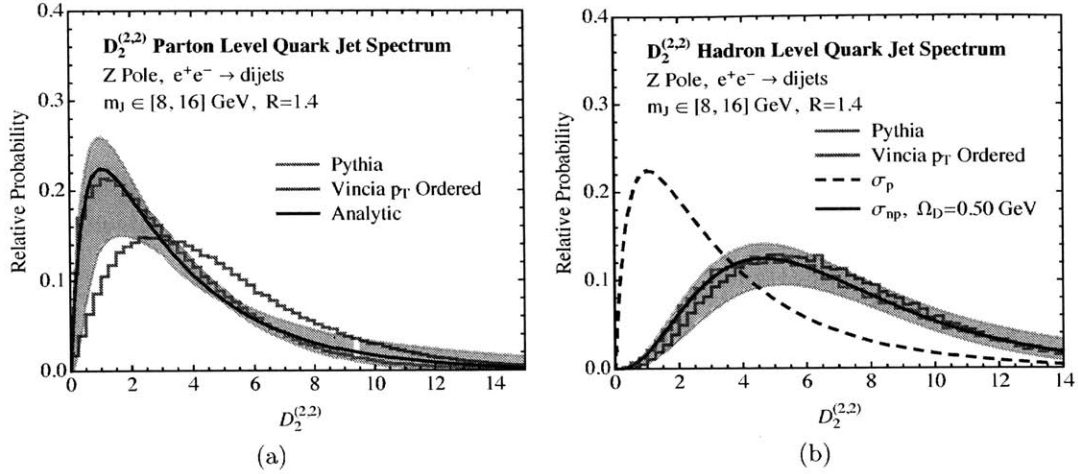
Figure 4-29: A comparison of the D_2 spectrum as measured on quark initiated jets at the Z pole from the PYTHIA and p_T-ordered VINCIA Monte Carlo generators to our analytic predictions. Results are shown both for parton level Monte Carlo compared with perturbative analytics in a), and for hadron level Monte Carlo compared with non-perturbative analytics in b).

small. The amount by which it is expected to decrease depends on another non-perturbative parameter, but is estimated in App. A.8 that Ω_D should decrease by approximately 0.1 GeV between our predictions at 1 TeV and those at LEP energies. This is an important consistency check on our results, but due to the large uncertainty, we cannot claim to probe this running over the scales that we have considered. The analytic perturbative spectrum is also shown for reference. Good overall agreement with both Monte Carlo generators is observed, and the discrepancy between the PYTHIA and VINCIA generators which was present at parton level is reduced, although still non-negligible. As was discussed in Sec. 4.5.5, it could also be compensated for by a modification of the non-perturbative shape parameter. In particular, the effect of hadronization is well captured by non-perturbative shape function. Hadronization has a significantly larger effect on the D_2 observable at Z pole energies than at 1 TeV. This demonstrates the consistency of our implementation of the non-perturbative corrections through the shape function, which predicts the scaling of the shift in the first moment through Eq. (4.80).

Unlike for the D_2 distributions at 1 TeV, where the effect of hadronization was well de-

206

scribed only by a shift in the first moment, at LEP energies the hadronization also has a non-trivial effect on the shape of the distribution. This can clearly be seen by comparing the dashed perturbative spectrum and the non-perturbative results in Fig. 4-29b. While our factorization of non-perturbative effects in terms of a shape function is completely generic, it is only the first moment of the shape function which is universal, with the full non-perturbative shape function being in general observable dependent. However, the modification in the shape of the D_2 spectrum due to hadronization effects seems to be quite well captured by the shape function of Eq. (4.62). In our plots we do not include any uncertainties due to the form of the non-perturbative shape function, despite the fact that they are the dominant effect throughout most of the hadronized distribution. More general shape functions, and a study of their associated uncertainties could be studied along the lines of Ref. [195], although this is beyond the scope of this chapter, and could only be justified if the perturbative components of our calculation were computed to a higher level of accuracy.

Since the D_2 spectrum is sensitive to the emissions from the gluon subjet, it is sensitive to the radiation pattern generated by a gluon, and could potentially be used to improve the Monte Carlo description of gluons and the modeling of color coherence effects. In contrast to most observables which have been used for tuning Monte Carlos to LEP data, such as the jet mass which is set by a single emission, D_2 requires two emissions off of the initiating quark to be non-zero, and therefore can be used as a more detailed probe of the perturbative shower. Although non-perturbative effects play a large role for jets in this energy range, we have shown that our factorization theorem allows us to cleanly separate perturbative from non-perturbative effects, which could be useful when tuning Monte Carlo generators, allowing one to disentangle genuine perturbative effects which should be well described by the Monte Carlo shower, from effects which should be captured by the hadronization model. We believe that higher order calculations of QCD jet shapes sensitive to three particle correlations, such as D_2, and their use in Monte Carlo tunings is therefore well motivated.

For reference, in App. A.9 we show a collection of $e_2^{(2)}$ distributions measured at the Z pole, at both parton and hadron level for both the VINCIA and PYTHIA event generators.

207

Unlike for the D_2 observable, the VINCIA and PYTHIA generators agree both at parton and hadron level to an excellent degree. This is of course expected due to the fact that these Monte Carlos have been tuned to LEP event shapes, but further emphasizes the fact that D_2, and other observables sensitive to additional emissions, provide a more detailed probe of the perturbative shower.

4.7 Looking Towards the LHC

Throughout this chapter, we have restricted our analysis to e^+e^- colliders so that we could ignore subtleties with initial state radiation, pile-up and other features important at hadron colliders. However, it is precisely for including these effects that a rigorous factorization based approach to jet substructure, such as that presented in this chapter, will prove most essential. In this section, we discuss the extension to the LHC and in particular to what extent conclusions for e^+e^- colliders holds for the LHC.

The energy correlation functions have a natural longitudinally-invariant generalization relevant for pp colliders, which is given by [30, 5]

$$
\begin{aligned}
e_2^{(\beta)} &= \frac{1}{p_{TJ}^2} \sum_{1 \le i < j \le n_J} p_{Ti} p_{Tj} R_{ij}^\beta \ , \\
e_3^{(\beta)} &= \frac{1}{p_{TJ}^3} \sum_{1 \le i < j < k \le n_J} p_{Ti} p_{Tj} p_{Tk} R_{ij}^\beta R_{ik}^\beta R_{jk}^\beta \ .
\end{aligned}
\tag{4.81}
$$

Here p_{TJ} is the transverse momentum of the jet with respect to the beam, p_{Ti} is the transverse momentum of particle i, and n_J is the number of particles contained in the jet. The boost-invariant angle $R_{ij}^2 = (\phi_i - \phi_j)^2 + (y_i - y_j)^2$ is defined as the Euclidean distance in the azimuth-rapidity plane. For central rapidity jets, which we will restrict ourselves to in this section, the power counting discussion of Sec. 4.2 is unmodified. Therefore, the same conclusions for the form of the optimal observable, D_2, as well as the range of angular exponents, apply. A simplified version of the $D_2^{(\alpha,\beta)}$ variable, restricted to have equal angular exponents $\alpha = \beta$, was used in Ref. [5], for jet substructure studies at the LHC.

It is in principle straightforward to extend the factorization theorems for D_2 to hadronic colliders, where D_2 is measured on a single jet in an exclusive N-jet event. Factorization theorems for exclusive N-jet production defined using N-jettiness [20, 104] or with a p_T-veto [208, 209] on additional radiation exist and could be combined with the factorization theorems of Sec. 4.3 to describe the jet substructure. We now briefly discuss how each of these factorization theorems can be interfaced with the presence of additional eikonal lines, representing either additional jets or beam directions in pp collisions.

Recall from Sec. 4.3.1, that the collinear subjets factorization theorem is formulated as a refactorization of the jet function for a particular jet in the n direction, and it is therefore insensitive to the global color structure of the event, seeing only the total color. Intuitively, the collinear-soft modes are boosted, and therefore all additional Wilson lines in the event are grouped in the $\bar{n}$ direction. Furthermore, the global soft modes, which resolve the global color structure of the event do not resolve the jet substructure. This property of the collinear subjets factorization theorem has the feature that it can be trivially combined with a factorization theorem with an arbitrary number of eikonal lines, without complicating the color structure. All that is then required, apart from the substructure components, is the addition of an additional measurement function in to the global soft function. Indeed, this extension has been discussed in detail in Ref. [105]. This same property is of course also true for the soft haze factorization theorem, as no additional Wilson lines are required to describe the jet substructure in the first place.

However, for the soft subjet factorization theorem, the presence of additional Wilson lines does significantly complicate the factorization from a calculational perspective. In particular, since the subjet is soft, arising from a refactorization of the soft function, it is emitted coherently from the N-eikonal line structure as a whole, requiring a proper treatment of all color correlations, which becomes complicated with even a few additional Wilson lines. A conjectural proposal for the all orders soft subjet factorization theorem with N-eikonal lines was given in Ref. [11], where the soft subjet factorization theorem was first proposed and studied in the large N_c limit. However, more work is required to understand its structure;
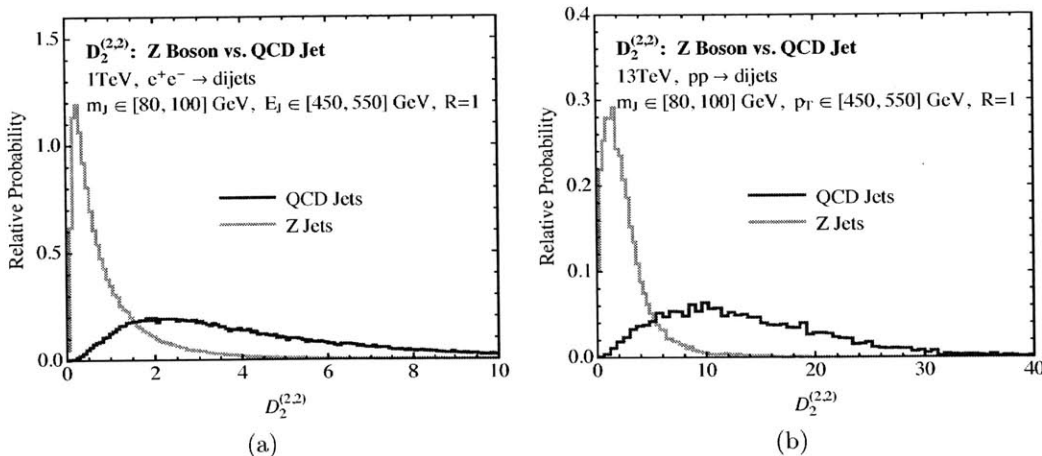
Figure 4-30: A comparison of the $D_2^{(2,2)}$ distributions for signal and background jets. a) Distributions for $R = 1$ jets at a 1 TeV e^+e^- collider. b) Distributions for $R = 1$ jets at the 13 TeV LHC, for jets with transverse momenta in the range $p_T \in [450, 550]$ GeV.

and an efficient organization of the color correlations at finite N_c. Furthermore for the soft subjet factorization theorem, the final soft function has an additional eikonal line, since the jet substructure is resolved by the long wavelength global soft modes, further complicating the calculation (although there has recently been some progress in the computation of soft functions [210, 211]). We emphasize however, that these are purely technical complications, and believe that the extension to a calculation of jet substructure in pp would be well worthwhile for improving our understanding of analytic jet substructure. Furthermore, depending on the relevant boosts and jet radii, the techniques of this chapter could be used to identify whether the soft subjet factorization theorem plays an important role, or could be formally neglected, simplifying the calculation in more complicated cases.

For these reasons, a full calculation in pp is well beyond the current scope of this initial investigation. We will instead restrict ourselves to a brief Monte Carlo study comparing the properties of D_2 in e^+e^- and pp to show that the distributions exhibit similar features. In Fig. 4-30 we compare the Monte Carlo predictions for $D_2^{(2,2)}$ as measured in e^+e^- collisions, shown in Fig. 4-30a, and pp collisions, shown in Fig. 4-30b. For e^+e^- collisions, the event selection is identical to earlier. For pp collisions, we generate background events from the

parton-level process $pp \to q\bar{q}$ and signal events from $pp \to ZZ \to q\bar{q}q\bar{q}$ events, where q denotes a massless quark, with PYTHIA 8.205 at the 13 TeV LHC.[25] Jets are clustered with the anti-k_T algorithm with radius $R = 1.0$, and using the WTA recombination scheme, with a p_T metric. We cut on the transverse momentum of the hardest jet, requiring $p_T \in [450, 550]$ GeV, and on the jet mass requiring $m_J \in [80, 100]$ GeV. These are chosen to be similar to the cuts on the jets for the case of e^+e^-, although they are of course not identical, and strict comparisons should not be made between the two cases. The shapes and general features of the D_2 distributions at the two colliders are very similar. There is a relative scaling between the D_2 distributions in e^+e^- and pp due to the different observable definitions. The e^+e^- definition uses the $1 - \cos(\theta_{ij})$ measure of Eq. (4.1), while the pp definition uses the boost invariant definition in terms of R_{ij}, as in Eq. (4.81). Since the $e_3^{(\alpha)}$ observable correlates particles of separation up to $2R$, where R is the jet radius, for $\alpha = \beta = 2$, this gives an expected factor of 4 difference between the two cases, as is approximately observed in Fig. 4-30.

The similar behavior of the e^+e^- and pp distributions suggests that a complete a calculation using our techniques would provide an excellent description of the D_2 distribution at a hadron collider, as we have found for e^+e^-. Such a calculation would also be interesting to better understand the effects of initial state radiation on the D_2 distribution. A simple setting where this calculation would be feasible, for example, would be to consider measuring the D_2 distribution on a jet recoiling against a color-singlet such as a W, Z or H boson, as was used in Ref. [84] to perform a NNLL calculation of the jet mass. Although the effects of non-global logarithms would need to be understood, and could play an important role, recent progress in this area suggests that this issue could be addressed, either by direct resummation of the NGLs [212, 213, 214, 215, 11], or through the use of jet grooming algorithms

---

[25]Since we only briefly mention the case of pp colliders, we do not perform a systematic study of the variation of the D_2 distribution in pp with different Monte Carlo generators, as we did for the case of e^+e^-. However, we believe that this is essential in any jet substructure study at pp, as we expect large variations will be present, as in the e^+e^- case. It would be particularly interesting to compare a p_T-ordered dipole-antenna shower, such as was recently implemented for pp in DIRE [187], with the PYTHIA and HERWIG++ generators which are more commonly used in jet substructure studies at the LHC. As was found for the case of e^+e^-, we expect a p_T-ordered dipole-antenna shower to provide the most accurate description of substructure observables at a pp collider, in particular, those sensitive to wide angle soft radiation.

which remove NGLs [86, 87, 40]. While it is truly uncorrelated with the jet, the effect of radiation from pile-up on D_2 could also be mitigated using similar jet grooming algorithms.

4.8 Conclusions

In this chapter we have presented a novel approach to the factorization of jet substructure observables, and applied it to the identification of two-prong substructure. Instead of starting with a given two-prong discriminant, we used the energy correlation functions as a basis of IRC safe observables to isolate the possible subjet configurations. We then studied the phase space defined by these IRC safe observables and proved all orders factorization theorems in each region of phase space. This procedure naturally identified an observable, D_2, which we argued provided optimal discrimination power, and which preserved the factorization properties of the individual factorization theorems describing different regions of the phase space defined by our basis of observables. We showed that a factorized description of this observable could be obtained by merging the different factorization theorems, and introduced a novel zero bin procedure in factorization theorem space to implement this merging. An important benefit of this approach is that our factorization theorems are valid to all orders in α_s at leading power and therefore provide a systematically improvable description of D_2.

Using our factorized description of the D_2 observable, we presented a numerical study of our results at an e^+e^- collider, for both the signal and background distributions, resulting in analytic boosted Z boson versus massive QCD jet discrimination predictions. We compared with a variety of Monte Carlo generators, and demonstrated that the low D_2 region, where a hard two-prong substructure is resolved, is a sensitive probe of the Monte Carlo parton shower description. We also studied the effect of non-perturbative corrections, showing that they can be well-described using a simple shape function, and related the single parameter of this shape function to a universal non-perturbative matrix element measured at LEP. This is vital for comparing our calculation with data.

Because our calculation presents the first factorized description of a two-prong discriminant jet observable in both signal and background regions, there are a large number of

directions for future study which are of great interest. First, our calculation was presented in the context of jets produced in e^+e^- collisions. For applications at the LHC, where jet substructure plays a vital role, it is important to extend the calculation to jets produced at a pp collider. The factorization theorem we presented straightforwardly generalizes to pp colliders with only complications due to soft radiation from the beams and the more complicated color structure of the hard interaction. The treatment of both these effects are well-understood and their inclusion in a jet substructure calculation would allow the first precision comparisons of calculations with data.

An interesting potential application of our factorization theorems, and merging procedures, which describe in a more differential way the substructure of jets, is to improve jet shape based subtraction schemes for QCD calculations at NNLO and beyond. Quite recently, subtractions based on the N-jettiness observable [104] have been used to perform NNLO calculations in QCD [216, 217, 218]. This allowed, in particular, the calculation of W, $H + 1$ jet at NNLO [216, 217] ($H + 1$ jet at NNLO was also calculated using more traditional subtraction techniques in [219]). The use of more differential subtractions based on more differential factorization theorems would allow for more local, and potentially numerically more efficient subtractions.

It would also be interesting to apply our calculation approach to other observables. For example, the N-subjettiness observables [32, 31] are used extensively in jet substructure studies at the LHC, and it would be of significant phenomenological relevance to obtain a factorized description of these observables. The approach presented in this chapter could also be extended to study more differential observables, such as those used for boosted top discrimination, which can resolve three subjets. A generalization of the $D_2^{(\alpha,\beta)}$ observable, $D_3^{(\alpha,\beta,\gamma)}$, which resolves three prong structure was introduced in Ref. [8] (see also Ref. [220] where it was used for boosted top discrimination at a 100 TeV collider). The $D_3^{(\alpha,\beta,\gamma)}$ observable should exhibit similar factorization properties to that of $D_2^{(\alpha,\beta)}$, and hence should be calculable with similar techniques. A rigorous factorization will also prove essential in this case, allowing for the separation of perturbative and non-perturbative physics, as well as

213

effects associated with the finite top width [221, 168]. More generally, we anticipate that the approach to the factorization of jet substructure observables presented in this chapter will allow for the construction of more powerful jet substructure discriminants and will enable a more detailed analytic understanding of the substructure of high energy QCD jets.

Chapter 5

Employing Helicity Amplitudes for Resummation

5.1   Introduction

The production of hadronic jets is one of the most basic processes at particle colliders. Processes including a vector boson (W, Z, γ) or Higgs boson together with jets provide probes of the Standard Model (SM), and are also dominant backgrounds for many new-physics searches. Optimizing the precision and discovery potential of these channels requires accurate predictions of the SM backgrounds. Furthermore, the growth of the jet substructure field has sparked a renewed interest in the study of jets themselves, both for an improved understanding of QCD, and for applications to identify boosted heavy objects in and beyond the SM.

Precise predictions for jet production require perturbative calculations including both fixed-order corrections as well as logarithmic resummation. QCD corrections to processes with jets are typically enhanced due to phase space restrictions. Such restrictions often introduce sensitivity to low momentum scales, p, of order a few tens of GeV, in addition to the hard scale, Q, which is of order the partonic center-of-mass energy. In this case, the perturbative series contains large double logarithms $\alpha_s^n \ln^m(p/Q)$ with $m \leq 2n$. To obtain

the best possible perturbative predictions, these logarithms should be resummed to all orders in α_s.

There has been tremendous progress in the calculation of fixed-order perturbative amplitudes in QCD using the spinor helicity formalism [222, 223, 224, 225], color ordering techniques [226, 227, 228, 229] and unitarity based methods [230, 231]. NLO predictions are now available for a large number of high multiplicity final states, including $pp \to V+$ up to 5 jets [232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242], $pp \to$ up to 5 jets [243, 244, 245, 246, 173, 247, 248, 249], and $pp \to H+$ up to 3 jets [250, 251, 252, 253, 254, 255, 256, 257, 258], and there are many efforts [259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 107, 275] to fully automatize the computation of one-loop corrections to generic helicity amplitudes.

For high-multiplicity jet events, the resummation of large logarithms is typically achieved with parton shower Monte Carlo programs. Here, the hard process enters through tree-level (and also one-loop) matrix elements and the QCD corrections due to final-state and initial-state radiation are described by the parton shower. The parton shower resums logarithms at the leading logarithmic (LL) accuracy, with some subleading improvements, but it is difficult to reliably assess and systematically improve its logarithmic accuracy.

The approach we will take in this chapter is to match onto soft-collinear effective theory (SCET) [21, 22, 23, 24], the effective theory describing the soft and collinear limits of QCD. In SCET, the QCD corrections at the hard scale are captured by process-dependent Wilson coefficients. The low-energy QCD dynamics does not depend on the details of the hard scattering (other than the underlying Born kinematics), similar to the parton shower picture. Resummation in SCET is achieved analytically through renormalization group evolution (RGE) in the effective theory, allowing one to systematically improve the logarithmic accuracy and assess the associated perturbative uncertainties. For example, for dijet event shape variables in e^+e^- collisions, SCET has enabled resummation to N^{3}LL accuracy and global fits for $\alpha_s(m_Z)$ [276, 81, 2, 206, 1, 207]. The analytic higher-order resummation can also be used to improve the Monte Carlo parton-shower description [277, 278, 279]. Further-

216

more, SCET allows for the direct calculation of exclusive jet cross sections, eliminating the need for numerical subtraction schemes for real emissions up to power corrections.

An important prerequisite for employing SCET is to obtain the hard matching coefficients, which are extracted from the fixed-order QCD amplitudes. The matching for $V + 2$ parton and $H + 2$ parton processes is well known from the QCD quark and gluon form factors, and is known to three loops [280, 281, 2]. The matching for $V + 3$ partons [282, 283, 284, 285], and $H + 3$ partons [208, 84, 286, 287], has been performed at both NLO and NNLO. Partonic processes with four external quarks have been studied in SCET in Refs. [288, 289, 290, 291, 292, 293, 294, 295], and the matching for all massless $2 \to 2$ processes has been obtained at NLO in Ref. [296] and recently at NNLO in Ref. [297].

For high-multiplicity processes, the usual approach to constructing an operator basis with explicit Lorentz indices and gamma matrices is laborious. In this chapter, we introduce a convenient formalism, based on helicity operators, which allows for a seamless matching for higher multiplicity processes onto SCET. A first look at the formalism discussed here was already given in Ref. [298]. Indeed, results for helicity amplitudes are already employed in the SCET matching calculations mentioned above, though without the construction of corresponding SCET operators.

In the spinor helicity formalism, the individual helicity amplitudes (i.e. the amplitudes for given fixed external helicities) are calculated, as opposed to calculating the amplitude for arbitrary external spins in one step and then summing over all spins at the end. One advantage is that the individual helicity amplitudes typically yield more compact expressions. And since they correspond to distinct external states, they can be squared and summed at the end. Helicity amplitudes remove the large redundancies in the usual description of (external) gauge fields, allowing for much simplified calculations particularly for amplitudes with many external gluons.

As we will see, this helicity-based approach is also advantageous in SCET. In SCET, as we will review in Sec. 5.2.2, collinear fields carry label directions corresponding to the directions of jets in the process, which provide natural lightlike vectors with which to de-

fine fields of definite helicity. As we will demonstrate, the construction of an appropriate operator basis becomes simple when using operators built out of fields with definite helicity. Furthermore, using such a helicity operator basis greatly facilitates the matching of QCD onto SCET, because one can directly utilize the known QCD helicity amplitudes for the matching. Together, this substantially simplifies the study of high-multiplicity jet processes with SCET.

5.1.1 Overview

Consider a process with N final-state jets and L leptons, photons, or other nonstrongly interacting particles, with the underlying hard Born process

$$\kappa_a(q_a)\,\kappa_b(q_b) \to \kappa_1(q_1) \cdots \kappa_{N+L}(q_{N+L})\,, \tag{5.1}$$

where $\kappa_{a,b}$ denote the colliding partons, and κ_i denote the outgoing quarks, gluons, leptons, and other particles with momenta q_i. The incoming partons are along the beam directions, $q_{a,b}^\mu = x_{a,b} P_{a,b}^\mu$, where $x_{a,b}$ are the momentum fractions and $P_{a,b}^\mu$ the (anti)proton momenta. For definiteness, we consider two colliding partons, but our discussion of the matching will be completely crossing symmetric, so it applies equally well to ep and ee collisions.

In SCET, the active-parton exclusive jet cross section corresponding to Eq. (5.1) can be proven to factorize for a variety of jet resolution variables.[1] The factorized expression for the exclusive jet cross section can be written schematically in the form

$$d\sigma = \int dx_a\, dx_b\, d\Phi_{N+L}(q_a + q_b; q_1, \ldots)\, M(\{q_i\}) \tag{5.2}$$
$$\times \sum_\kappa \mathrm{tr}\left[\widehat{H}_\kappa(\{q_i\})\widehat{S}_\kappa\right] \otimes \left[B_{\kappa_a} B_{\kappa_b} \prod_J J_{\kappa_J}\right] + \cdots.$$

Here, $d\Phi_{N+L}(q_a + q_b; q_1, \ldots)$ denotes the Lorentz-invariant phase space for the Born process

---

[1]Here active parton refers to initial-state quarks or gluons. Proofs of factorization with initial-state hadrons must also account for effects due to Glaubers [16], which may or may not cancel, and whose relevance depends on the observable in question [17, 18].

in Eq. (5.1), and $M(\{q_i\})$ denotes the measurement made on the hard momenta of the jets (which in the factorization are approximated by the Born momenta q_i). The dependence on the underlying hard interaction is encoded in the hard function $\widehat{H}_\kappa(\{q_i\})$, where $\{q_i\} \equiv \{q_1, \ldots, q_{N+L}\}$, the sum over $\kappa \equiv \{\kappa_a, \kappa_b, \ldots \kappa_{N+L}\}$ is over all relevant partonic processes, and the trace is over color. Any dependence probing softer momenta, such as measuring jet masses or low p_Ts, as well as the choice of jet algorithm, will affect the precise form of the factorization, but not the hard function $\widehat{H}_\kappa$. This dependence enters through the definition of the soft function $\widehat{S}_\kappa$ (describing soft radiation), jet functions J_{κ_J} (describing energetic final-state radiation in the jets) and the beam functions B_i (describing energetic initial-state radiation along the beam direction). More precisely, the beam function is given by $B_i = \sum_{i'} \mathcal{I}_{ii'} \otimes f_{i'}$ with f_i the parton distributions of the incoming protons, and $\mathcal{I}_{ii'}$ a perturbatively calculable matching coefficient depending on the measurement definition [20]. The ellipses at the end of Eq. (5.2) denote power-suppressed corrections. All functions in the factorized cross section depend only on the physics at a single scale. This allows one to evaluate all functions at their own natural scale, and then evolve them to a common scale using their RGE. This procedure resums the large logarithms of scale ratios appearing in the cross section to all orders in perturbation theory.

The explicit form of the factorization theorem in Eq. (5.2), including field-theoretic definitions for the jet, beam, and soft functions is known for a number of exclusive jet cross sections and measurements of interest. For example, factorization theorems exist for the N-jet cross section defined using N-jettiness [20, 104, 299, 300, 301, 84, 302, 303, 304]. These have also been utilized to include higher-order resummation in Monte Carlo programs [277, 278, 279], and are the basis of the N-jettiness subtraction method for fixed-order calculations [217, 218]. In addition, there has been a focus on color-singlet production at small q_T [305, 306, 307, 308, 309], as well as the factorization of processes defined with jet algorithms [310, 311, 167, 90, 312, 313, 314, 191, 315, 208, 316, 13, 317, 318, 319, 320, 321], jet shape variables [203, 322, 105, 85, 82, 10, 37, 138, 34, 323, 11, 12], or fragmentation properties [324, 325, 326, 327, 141, 142, 328, 329, 330] for identified jets. The same hard

functions also appear in threshold resummation factorization formulas, which are often used to obtain an approximate higher order result for inclusive cross sections.

The focus of our chapter is the hard function $\widehat{H}_\kappa(\{q_i\})$ in Eq. (5.2), which contains the process-dependent underlying hard interaction of Eq. (5.1), but is independent of the particular measurement. In SCET, the dependence on the hard interaction is encoded in the Wilson coefficients, $\vec{C}$, of a basis of operators built out of SCET fields. The Wilson coefficients can be calculated through a matching calculation from QCD onto the effective theory. The hard function appearing in the factorization theorem is then given by

$$\widehat{H}_\kappa(\{q_i\}) = \sum_{\{\lambda_i\}} \vec{C}_{\lambda_1 \cdots (\cdots \lambda_n)}(\{q_i\}) \, \vec{C}^\dagger_{\lambda_1 \cdots (\cdots \lambda_n)}(\{q_i\}) \,. \tag{5.3}$$

Here, the $\{\lambda_i\}$ denote helicity labels and the sum runs over all relevant helicity configurations. The $\vec{C}$ are vectors in color space, and the hard function is therefore a matrix in color space.

For processes of higher multiplicities, the construction of a complete basis of SCET operators, and the subsequent matching calculation, becomes laborious due to the proliferation of Lorentz and color structures, similar to the case of high-multiplicity fixed-order calculations using standard Feynman diagrams. The use of SCET helicity fields introduced in this chapter, combined with analogous color management techniques as used in the calculation of amplitudes, makes the construction of an operator basis extremely simple, even in the case of high-multiplicity processes. Furthermore, with this basis choice, the SCET Wilson coefficients are precisely given by the IR-finite parts of the color-ordered QCD helicity amplitudes, rendering the matching procedure almost trivial. Combining the results for the hard function with known results for the soft, jet, and beam functions, then allows for the resummation of jet observables in higher multiplicity processes, which are ubiquitous at the LHC.

The remainder of this chapter is organized as follows. In Sec. 5.2.1, we review the notation for the spinor-helicity formalism. Additional useful helicity and color identities can be found in Appendix B.1. We provide a brief summary of SCET in Sec. 5.2.2. In Sec. 5.3, we introduce SCET helicity fields and operators, and describe the construction of the helicity

and color basis, as well as its symmetry properties. In Sec. 5.4, we discuss the matching from QCD onto the SCET helicity operators, including a discussion of the dependence on the regularization and renormalization scheme. We then demonstrate the matching explicitly for $H + 0, 1, 2$ jets in Sec. 5.5, $V + 0, 1, 2$ jets in Sec. 5.6, and $pp \to 2, 3$ jets in Sec. 5.7. Explicit results for the required helicity amplitudes are collected in the appendices. In Sec. 5.8, we discuss the general renormalization group evolution of the hard coefficients, which involves mixing between different color structures, to all orders. We give explicit results for the anomalous dimensions for up to 4 colored particles plus an arbitrary number of uncolored particles. We conclude in Sec. 5.9.

5.2   Notation

5.2.1   Helicity Formalism

We will use the standard notation for the spinor algebra (for a review see for example Refs. [331, 332]). Consider the four-component spinor $u(p)$ of a massless Dirac particle with momentum p, satisfying the massless Dirac equation,

$$\slashed{p}\, u(p) = 0 , \qquad p^2 = 0 . \tag{5.4}$$

The charge conjugate (antiparticle) spinor $v(p)$ also satisfies Eq. (5.4), and we can choose a representation such that $v(p) = u(p)$. The spinors and conjugate spinors for the two helicity states are denoted by

$$|p\pm\rangle = \frac{1 \pm \gamma_5}{2}\, u(p) ,$$
$$\langle p\pm| = \mathrm{sgn}(p^0)\, \bar{u}(p)\, \frac{1 \mp \gamma_5}{2} . \tag{5.5}$$

For massless particles chirality and helicity agree while for antiparticles they are opposite, so $|p+\rangle = u_+(p) = v_-(p)$ corresponds to positive (negative) helicity for particles (antiparticles).

The spinors $|p\pm\rangle$ are defined by Eqs. (5.4) and (5.5) for both physical ($p^0 > 0$) and unphysical ($p^0 < 0$) momenta. Their explicit expression, including our overall phase convention, is given in Appendix B.1.1.

The spinor products are denoted by

$$\langle pq \rangle = \langle p-\rangle\, q+ \,, \qquad [pq] = \langle p+\rangle\, q- \,. \tag{5.6}$$

They satisfy

$$\langle pq \rangle = -\langle qp \rangle \,, \quad [pq] = -[qp] \,, \quad \langle pq \rangle [qp] = 2p \cdot q \,. \tag{5.7}$$

Additional relations are collected in Appendix B.1.1. The minus sign for $p^0 < 0$ in Eq. (5.5) is included so the spinor relations are invariant under inverting the signs of momenta, $p \to -p$, when crossing particles between the initial and final state, e.g. $\langle (-p)q \rangle [q(-p)] = 2(-p) \cdot q$.

If there are several momenta p_i, it is common to abbreviate

$$|p_i\pm\rangle = |i\pm\rangle \,, \qquad \langle p_i p_j \rangle = \langle ij \rangle \,, \qquad [p_i p_j] = [ij] \,. \tag{5.8}$$

The polarization vectors of an outgoing gluon with momentum p are given in the helicity formalism by

$$\varepsilon_+^\mu(p,k) = \frac{\langle p+|\gamma^\mu|k+\rangle}{\sqrt{2}\langle kp \rangle} \,, \quad \varepsilon_-^\mu(p,k) = -\frac{\langle p-|\gamma^\mu|k-\rangle}{\sqrt{2}[kp]} \,, \tag{5.9}$$

where k is an arbitrary reference vector with $k^2 = 0$, which fixes the gauge of the external gluons. Using the relations in Appendix B.1.1, it is easy to check that

$$p \cdot \varepsilon_\pm(p,k) = k \cdot \varepsilon_\pm(p,k) = 0 \,,$$
$$\varepsilon_\pm(p,k) \cdot \varepsilon_\pm(p,k) = 0 \,,$$
$$\varepsilon_\pm(p,k) \cdot \varepsilon_\mp(p,k) = -1 \,,$$
$$\varepsilon_\pm^*(p,k) = \varepsilon_\mp(p,k) \,, \tag{5.10}$$

as is required for physical polarization vectors. With $p^\mu = E(1,0,0,1)$, the choice $k^\mu = E(1,0,0,-1)$ yields the conventional

$$\varepsilon^\mu_\pm(p,k) = \frac{1}{\sqrt{2}}(0,1,\mp i,0)\,. \tag{5.11}$$

5.2.2 SCET

We now briefly review the SCET concepts and notation that will be required for this section. As has been discussed, Soft-collinear effective theory is an effective field theory of QCD that describes the interactions of collinear and soft particles [21, 22, 23, 24] in the presence of a hard interaction.[2] Collinear particles are characterized by having large energy and small invariant mass. To separate the large and small momentum components, it is convenient to use light-cone coordinates. We define two light-cone vectors

$$n^\mu = (1,\vec{n})\,, \qquad \bar{n}^\mu = (1,-\vec{n})\,, \tag{5.12}$$

with $\vec{n}$ a unit three-vector, which satisfy $n^2 = \bar{n}^2 = 0$ and $n \cdot \bar{n} = 2$. Any four-momentum p can be decomposed as

$$p^\mu = \bar{n}\cdot p\,\frac{n^\mu}{2} + n\cdot p\,\frac{\bar{n}^\mu}{2} + p^\mu_{n\perp}\,. \tag{5.13}$$

An "n-collinear" particle has momentum p close to the $\vec{n}$ direction, so that p scales as $(n\cdot p, \bar{n}\cdot p, p_{n\perp}) \sim \bar{n}\cdot p\ (\lambda^2, 1, \lambda)$, with $\lambda \ll 1$ a small parameter. For example, for a jet of collinear particles in the $\vec{n}$ direction with total momentum p_J, $\bar{n}\cdot p_J \simeq 2E_J$ corresponds to the large energy of the jet, while $n\cdot p_J \simeq m_J^2/E_J \ll E_J$, where m_J is the jet mass, so $\lambda^2 \simeq m_J^2/E_J^2 \ll 1$.

To construct the fields of the effective theory, the momentum of n-collinear particles is

---

[2]Throughout this chapter, we will for simplicity use the notation of SCET$_\text{I}$. The theory SCET$_\text{II}$ [28] is required for a certain class of observables, for example p_T-dependent measurements or vetoes. The helicity operator formalism presented here applies identically to constructing SCET$_\text{II}$ operators. The collinear operators and matching coefficients are the same for both cases.

written as

$$p^\mu = \tilde{p}^\mu + k^\mu = \bar{n}\cdot\tilde{p}\,\frac{n^\mu}{2} + \tilde{p}_{n\perp}^\mu + k^\mu \,, \tag{5.14}$$

where $\bar{n}\cdot\tilde{p} \sim Q$ and $\tilde{p}_{n\perp} \sim \lambda Q$ are the large momentum components, while $k \sim \lambda^2 Q$ is a small residual momentum. Here, Q is the scale of the hard interaction, and the effective theory expansion is in powers of λ.

The SCET fields for n-collinear quarks and gluons, $\xi_{n,\tilde{p}}(x)$ and $A_{n,\tilde{p}}(x)$, are labeled by the collinear direction n and their large momentum $\tilde{p}$. They are written in position space with respect to the residual momentum and in momentum space with respect to the large momentum components. Derivatives acting on the fields pick out the residual momentum dependence, $i\partial^\mu \sim k \sim \lambda^2 Q$. The large label momentum is obtained from the label momentum operator $\mathcal{P}_n^\mu$, e.g. $\mathcal{P}_n^\mu \xi_{n,\tilde{p}} = \tilde{p}^\mu \xi_{n,\tilde{p}}$. If there are several fields, $\mathcal{P}_n$ returns the sum of the label momenta of all n-collinear fields. For convenience, we define $\overline{\mathcal{P}}_n = \bar{n}\cdot\mathcal{P}_n$, which picks out the large momentum component. Frequently, we will only keep the label n denoting the collinear direction, while the momentum labels are summed over (subject to momentum conservation) and are suppressed in our notation.

Collinear operators are constructed out of products of fields and Wilson lines that are invariant under collinear gauge transformations [22, 23]. The smallest building blocks are collinearly gauge-invariant quark and gluon fields, defined as

$$\chi_{n,\omega}(x) = \left[\delta(\omega - \overline{\mathcal{P}}_n)\, W_n^\dagger(x)\, \xi_n(x)\right],$$
$$\mathcal{B}_{n,\omega\perp}^\mu(x) = \frac{1}{g}\left[\delta(\omega + \overline{\mathcal{P}}_n)\, W_n^\dagger(x)\, iD_{n\perp}^\mu W_n(x)\right]. \tag{5.15}$$

With this definition of $\chi_{n,\omega}$, we have $\omega > 0$ for an incoming quark and $\omega < 0$ for an outgoing antiquark. For $\mathcal{B}_{n,\omega\perp}$, $\omega > 0$ ($\omega < 0$) corresponds to an outgoing (incoming) gluon. In Eq. (5.15)

$$iD_{n\perp}^\mu = \mathcal{P}_{n\perp}^\mu + gA_{n\perp}^\mu \,, \tag{5.16}$$

224

is the collinear covariant derivative and

$$W_n(x) = \left[\sum_{\text{perms}} \exp\left(-\frac{g}{\overline{\mathcal{P}}_n} \, \bar{n} \cdot A_n(x) \right) \right] \tag{5.17}$$

is a Wilson line of n-collinear gluons in label momentum space. The label operators $\overline{\mathcal{P}}_n$ in Eqs. (5.15) and (5.17) only act inside the square brackets. $W_n(x)$ sums up arbitrary emissions of n-collinear gluons from an $\bar{n}$-collinear quark or gluon, which are $\mathcal{O}(1)$ in the power counting. Since $W_n(x)$ is localized with respect to the residual position x, we can treat $\chi_{n,\omega}(x)$ and $\mathcal{B}_{n,\omega}^\mu(x)$ like local quark and gluon fields. For later use, we give the expansion of the collinear gluon field

$$\mathcal{B}_{n,\perp}^\mu = A_{n\perp}^\mu - \frac{p_\perp^\mu}{\bar{n} \cdot p} \bar{n} \cdot A_{n,p} + \cdots . \tag{5.18}$$

Here the ellipses denote terms in the expansion with more than 2 collinear gluon fields, which are not required for our matching calculations.

In our case the effective theory contains several collinear sectors, $n_1, n_2, \ldots$ [25], where the collinear fields for a given sector $n_i^\mu = (1, \vec{n}_i)$ describe a jet in the direction $\vec{n}_i$, and we also define $\bar{n}_i^\mu = (1, -\vec{n}_i)$. A fixed-order QCD amplitude with N colored legs is then matched onto operators in SCET with N different collinear fields. The different collinear directions have to be well separated, which means

$$n_i \cdot n_j \gg \lambda^2 \qquad \text{for} \qquad i \neq j . \tag{5.19}$$

The infrared singularities associated with collinear or soft limits of legs in QCD are entirely described by the Lagrangian and dynamics of SCET itself, so the QCD amplitudes are only used to describe the hard kinematics away from infrared singular limits.

Two different n_i and n_i' with $n_i \cdot n_i' \sim \lambda^2$ both describe the same jet and corresponding collinear physics. Thus, each collinear sector can be labeled by any member of a set of equivalent vectors, $\{n_i\}$, which are related by reparametrization invariance [26]. The simplest

way to perform the matching is to choose n_i such that the large label momentum is given by

$$\tilde{p}_i^\mu = \omega_i \, \frac{n_i^\mu}{2} \,, \tag{5.20}$$

with $\tilde{p}_{n_i\perp}^\mu = 0$.

In general, operators will have sums over distinct equivalence classes, $\{n_i\}$, and matrix elements select a representative vector to describe particles in a particular collinear direction. For many leading power applications there is only a single collinear field in each sector, and we may simply set the large label momentum of that building block field to that of the external parton using the following simple relation,

$$\int d\tilde{p} \, \tilde{\delta}(\tilde{p} - p) \, f(\tilde{p}) = f\left(\bar{n}_i \cdot p \, \frac{n_i}{2}\right) \,, \tag{5.21}$$

where p is collinear with the i'th jet. Here the tildes on the integration measure and delta function ensure that the integration over equivalence classes is properly implemented.[3] Because of this, at leading power, the issue of equivalence classes can largely be ignored.

Particles that exchange large momentum of $\mathcal{O}(Q)$ between different jets are off shell by $\mathcal{O}(n_i \cdot n_j \, Q^2)$. They are integrated out by matching QCD onto SCET. Before and after the hard interaction the jets described by the different collinear sectors evolve independently from each other, with only soft radiation between the jets. The corresponding soft degrees of freedom are described in the effective theory by soft quark and gluon fields, $q_s(x)$ and $A_s(x)$, which only have residual soft momentum dependence $i\partial^\mu \sim \lambda^2 Q$. They couple to the

---

[3]The precise definition of this delta function and measure are

$$\tilde{\delta}(\tilde{p}_i - p) \equiv \delta_{\{n_i\},p} \, \delta(\omega_i - \bar{n}_i \cdot p) \,,$$
$$\int d\tilde{p} \equiv \sum_{\{n_i\}} \int d\omega_i \,, \tag{5.22}$$

where

$$\delta_{\{n_i\},p} = \begin{cases} 1 & n_i \cdot p = \mathcal{O}(\lambda^2) \,, \\ 0 & \text{otherwise} \,. \end{cases} \tag{5.23}$$

The Kronecker delta is nonzero if the collinear momentum p is in the $\{n_i\}$ equivalence class, i.e. p is close enough to be considered as collinear with the ith jet. The sum in the second line of Eq. (5.22) runs over the different equivalence classes.

collinear sectors via the soft covariant derivative

$$iD_s^\mu = i\partial^\mu + gA_s^\mu\,, \tag{5.24}$$

acting on the collinear fields. At leading power in λ, n-collinear particles only couple to the $n \cdot A_s$ component of soft gluons, so the leading-power n-collinear Lagrangian only depends on $n \cdot D_s$. For example, for n-collinear quarks [22, 23]

$$\mathcal{L}_n = \bar{\xi}_n \Big(in \cdot D_s + g\, n \cdot A_n + i\slashed{\mathcal{D}}_{n\perp} W_n \frac{1}{\overline{\mathcal{P}}_n} W_n^\dagger i\slashed{\mathcal{D}}_{n\perp} \Big) \frac{\slashed{\bar{n}}}{2} \xi_n\,. \tag{5.25}$$

The leading-power n-collinear Lagrangian for gluons is given in Ref. [24].

5.3 SCET Operator Basis

In this section, we describe in detail how to construct a basis of helicity and color operators in SCET, which greatly simplifies the construction of a complete operator basis and also facilitates the matching process. Usually, a basis of SCET operators obeying the symmetries of the problem is constructed from the fields $\chi_{n,\omega}$, $\mathcal{B}_{n,\omega\perp}^\mu$, as well as Lorentz and color structures. This process becomes quite laborious due to the large number of structures which appear for higher multiplicity processes, and the reduction to a minimal basis of operators quickly becomes nontrivial. Instead, we work with a basis of operators with definite helicity structure constructed from scalar SCET building blocks, which, as we will show, has several advantages. First, this simplifies the construction of the operator basis, because each independent helicity configuration gives rise to an independent helicity operator. In this way, we automatically obtain the minimal number of independent operators as far as their Lorentz structure is concerned. Second, operators with distinct helicity structures do not mix under renormalization group evolution, as will be discussed in detail in Sec. 5.8. The reason is that distinct jets can only exchange soft gluons in SCET, which at leading order in the power counting means they can transfer color but not spin [see Eq. (5.25)]. Therefore, the only

227

nontrivial aspect of the operator basis is the color degrees of freedom. The different color structures mix under renormalization group evolution, but their mixing only depends on the color representations and not on the specific helicity configuration.

5.3.1 Helicity Fields

We start by defining quark and gluon fields of definite helicity, out of which we can build operators with a definite helicity structure. To simplify our discussion we will take all momenta and polarization vectors as outgoing, and label all fields and operators by their outgoing helicity and momenta. Crossing symmetry, and crossing relations are discussed in Sec. 5.3.6.

We define a gluon field of definite helicity[4]

$$
\mathcal{B}_{i\pm}^{a} = -\varepsilon_{\mp\mu}(n_i, \bar{n}_i)\, \mathcal{B}_{n_i,\omega_i \perp_i}^{a\mu} \,,
\tag{5.26}
$$

where a is an adjoint color index. For $n_i^{\mu} = (1, 0, 0, 1)$, we have

$$
\varepsilon_{\pm}^{\mu}(n_i, \bar{n}_i) = \frac{1}{\sqrt{2}}\,(0, 1, \mp i, 0)\,,
\tag{5.27}
$$

in which case

$$
\mathcal{B}_{i\pm}^{a} = \frac{1}{\sqrt{2}}\big(\mathcal{B}_{n_i,\omega_i\,\perp_i}^{a,1} \pm i\mathcal{B}_{n_i,\omega_i\,\perp_i}^{a,2}\big)\,.
\tag{5.28}
$$

For an external gluon with outgoing polarization vector $\varepsilon(p, k)$ and outgoing momentum p in the n_i-collinear direction, the contraction with the field $\mathcal{B}_{i\pm}^{a}$ contributes

$$
-\varepsilon_{\mp\mu}(n_i, \bar{n}_i)\Big[\varepsilon_{\perp_i}^{\mu}(p, k) - \frac{p_{\perp_i}^{\mu}}{\bar{n}_i \cdot p}\,\bar{n}_i \cdot \varepsilon(p, k)\Big]\,,
\tag{5.29}
$$

where we have used the expansion of the collinear gluon field given in Eq. (5.18). Since $\varepsilon_{\mp}(n_i, \bar{n}_i)$ is perpendicular to both n_i and $\bar{n}_i$, we can drop the $\perp_i$ labels in brackets. A convenient choice for the reference vector is to take $k = \bar{n}_i$, for which the second term in

---

[4]The label $\pm$ on $\mathcal{B}_\pm$ refers to helicity and should not be confused with light-cone components.

brackets vanishes. Equation (5.29) then becomes

$$-\varepsilon_{\mp}(n_i, \bar{n}_i) \cdot \varepsilon(p, \bar{n}_i)\,, \tag{5.30}$$

which is equal to 0 or 1 depending on the helicity of $\varepsilon(p, \bar{n}_i)$. Adopting this choice, the tree-level Feynman rules for an outgoing gluon with polarization $\pm$ (so $\varepsilon = \varepsilon_{\pm}$), momentum p (with $p^0 > 0$), and color a are

$$\begin{aligned}
\langle g_{\pm}^a(p)|\mathcal{B}_{i\pm}^b|0\rangle &= \delta^{ab}\,\tilde{\delta}(\tilde{p}_i - p)\,, \\
\langle g_{\mp}^a(p)|\mathcal{B}_{i\pm}^b|0\rangle &= 0\,.
\end{aligned} \tag{5.31}$$

Note that $\mathcal{B}_{i\pm}^b = \mathcal{B}_{i\pm}^b(0)$, so we do not get a phase from the residual momentum. Similarly, for an incoming gluon with incoming polarization $\mp$ ($\varepsilon = \varepsilon_{\mp}$, so $\varepsilon^* = \varepsilon_{\pm}$), incoming momentum $-p$ (with $p^0 < 0$), and color a, we have

$$\begin{aligned}
\langle 0|\mathcal{B}_{i\pm}^b|g_{\mp}^a(-p)\rangle &= \delta^{ab}\,\tilde{\delta}(\tilde{p}_i - p)\,, \\
\langle 0|\mathcal{B}_{i\pm}^b|g_{\pm}^a(-p)\rangle &= 0\,.
\end{aligned} \tag{5.32}$$

We define quark fields with definite helicity[5] as

$$\chi_{i\pm}^{\alpha} = \frac{1 \pm \gamma_5}{2}\,\chi_{n_i,-\omega_i}^{\alpha}\,, \qquad \bar{\chi}_{i\pm}^{\bar{\alpha}} = \bar{\chi}_{n_i,\omega_i}^{\bar{\alpha}}\,\frac{1 \mp \gamma_5}{2}\,, \tag{5.33}$$

where α and $\bar{\alpha}$ are fundamental and antifundamental color indices respectively.

For external quarks with n_i-collinear momentum p, the fields contribute factors of the form

$$\frac{1 \pm \gamma_5}{2}\,\frac{\slashed{n}_i\slashed{\bar{n}}_i}{4}\,u(p) = \frac{\slashed{n}_i\slashed{\bar{n}}_i}{4}\,|p\pm\rangle = |p\pm\rangle_{n_i}\,, \tag{5.34}$$

---

[5]Technically speaking chirality, although we work in a limit where all external quarks can be treated as massless.

where in the last equality, we have defined a shorthand notation $|p\pm\rangle_{n_i}$ for the SCET projected spinor. The spinor $|p\pm\rangle_{n_i}$ is proportional to $|n\pm\rangle$; see Eq. (B.30).

The tree-level Feynman rules for incoming ($p^0 < 0$) and outgoing ($p^0 > 0$) quarks with helicity $+/-$ and color α are then given by

$$\langle 0|\chi_{i+}^{\beta}|q_{+}^{\bar{\alpha}}(-p)\rangle = \delta^{\beta\bar{\alpha}}\,\tilde{\delta}(\tilde{p}_i - p)\,|(-p_i)+\rangle_{n_i}\,,$$

$$\langle 0|\chi_{i-}^{\beta}|q_{-}^{\bar{\alpha}}(-p)\rangle = \delta^{\beta\bar{\alpha}}\,\tilde{\delta}(\tilde{p}_i - p)\,|(-p_i)-\rangle_{n_i}\,,$$

$$\langle q_{+}^{\alpha}(p)|\bar{\chi}_{i+}^{\bar{\beta}}|0\rangle = \delta^{\alpha\bar{\beta}}\,\tilde{\delta}(\tilde{p}_i - p)\,_{n_i}\langle p_i+|\,,$$

$$\langle q_{-}^{\alpha}(p)|\bar{\chi}_{i-}^{\bar{\beta}}|0\rangle = \delta^{\alpha\bar{\beta}}\,\tilde{\delta}(\tilde{p}_i - p)\,_{n_i}\langle p_i-|\,, \tag{5.35}$$

and similarly for antiquarks

$$\langle 0|\bar{\chi}_{i+}^{\bar{\beta}}|\bar{q}_{-}^{\alpha}(-p)\rangle = \delta^{\alpha\bar{\beta}}\,\tilde{\delta}(\tilde{p}_i - p)\,_{n_i}\langle(-p_i)+|\,,$$

$$\langle 0|\bar{\chi}_{i-}^{\bar{\beta}}|\bar{q}_{+}^{\alpha}(-p)\rangle = \delta^{\alpha\bar{\beta}}\,\tilde{\delta}(\tilde{p}_i - p)\,_{n_i}\langle(-p_i)-|\,,$$

$$\langle \bar{q}_{-}^{\bar{\alpha}}(p)|\chi_{i+}^{\beta}|0\rangle = \delta^{\beta\bar{\alpha}}\,\tilde{\delta}(\tilde{p}_i - p)\,|p_i+\rangle_{n_i}\,,$$

$$\langle \bar{q}_{+}^{\bar{\alpha}}(p)|\chi_{i-}^{\beta}|0\rangle = \delta^{\beta\bar{\alpha}}\,\tilde{\delta}(\tilde{p}_i - p)\,|p_i-\rangle_{n_i}\,. \tag{5.36}$$

The corresponding Feynman rules with the helicity of the external (anti)quark flipped vanish.

To avoid the explicit spinors in Eqs. (5.35) and (5.36), and exploit the fact that fermions come in pairs, we also define fermionic vector currents of definite helicity

$$J_{ij+}^{\bar{\alpha}\beta} = \frac{\sqrt{2}\,\varepsilon_{-}^{\mu}(n_i, n_j)}{\sqrt{\omega_i\,\omega_j}}\,\frac{\bar{\chi}_{i+}^{\bar{\alpha}}\,\gamma_\mu\chi_{j+}^{\beta}}{\langle n_i n_j\rangle}\,,$$

$$J_{ij-}^{\bar{\alpha}\beta} = -\frac{\sqrt{2}\,\varepsilon_{+}^{\mu}(n_i, n_j)}{\sqrt{\omega_i\,\omega_j}}\,\frac{\bar{\chi}_{i-}^{\bar{\alpha}}\,\gamma_\mu\chi_{j-}^{\beta}}{[n_i n_j]}\,, \tag{5.37}$$

where $\omega_i = \bar{n}_i \cdot \tilde{p}_i$ from Eq. (5.20), as well as a scalar current

$$J_{ij0}^{\bar{\alpha}\beta} = \frac{2}{\sqrt{\omega_i\,\omega_j}}\,\frac{\bar{\chi}_{i+}^{\bar{\alpha}}\chi_{j-}^{\beta}}{[n_i n_j]}\,,$$

$$(J^\dagger)_{ij0}^{\bar\alpha\beta} = \frac{2}{\sqrt{\omega_i\,\omega_j}} \frac{\bar\chi_{i-}^{\bar\alpha}\chi_{j+}^{\beta}}{\langle n_i n_j\rangle}. \tag{5.38}$$

In Eqs. (5.37) and (5.38) the flavor labels of the quarks have not been made explicit, but in general the two quark fields in a current can have different flavors (for example in W production). Since we are using a basis of physical polarization states it is not necessary to introduce more complicated Dirac structures. For example, pseudovector and pseudoscalar currents, which are usually introduced using γ^5, are incorporated through the relative coefficients of operators involving J_+, J_- or J_0, $J_0^\dagger$. As we shall see, this greatly simplifies the construction of the operator basis in the effective theory.

At leading power, there is a single collinear field in each collinear sector, so we can choose $n_i^\mu = p_i^\mu/p_i^0$ to represent the equivalence class $\{n_i\}$, so that $p_i^\mu = \frac{1}{2}\bar n \cdot p_i\, n_i^\mu$ which gives

$$|p\pm\rangle_{n_i} = |p\pm\rangle = \left|\bar n_i \cdot p\, \frac{n_i}{2}\pm\right\rangle = \sqrt{\frac{\bar n_i \cdot p}{2}}\,|n_i\pm\rangle. \tag{5.39}$$

Since we always work at leading power in this chapter, we will always make this choice to simplify the matching. With this choice, the tree-level Feynman rules for the fermion currents are

$$\begin{aligned}
&\left\langle q_+^{\alpha_1}(p_1)\,\bar q_-^{\bar\alpha_2}(p_2)\big| J_{12+}^{\bar\beta_1\beta_2}\big|0\right\rangle \\
&\quad = \delta^{\alpha_1\bar\beta_1}\,\delta^{\beta_2\bar\alpha_2}\,\tilde\delta(\tilde p_1 - p_1)\,\tilde\delta(\tilde p_2 - p_2), \\
&\left\langle q_-^{\alpha_1}(p_1)\,\bar q_+^{\bar\alpha_2}(p_2)\big| J_{12-}^{\bar\beta_1\beta_2}\big|0\right\rangle \\
&\quad = \delta^{\alpha_1\bar\beta_1}\,\delta^{\beta_2\bar\alpha_2}\,\tilde\delta(\tilde p_1 - p_1)\,\tilde\delta(\tilde p_2 - p_2), \\
&\left\langle q_+^{\alpha_1}(p_1)\,\bar q_+^{\bar\alpha_2}(p_2)\big| J_{120}^{\bar\beta_1\beta_2}\big|0\right\rangle \\
&\quad = \delta^{\alpha_1\bar\beta_1}\,\delta^{\beta_2\bar\alpha_2}\,\tilde\delta(\tilde p_1 - p_1)\,\tilde\delta(\tilde p_2 - p_2), \\
&\left\langle q_-^{\alpha_1}(p_1)\,\bar q_-^{\bar\alpha_2}(p_2)\big| (J^\dagger)_{120}^{\bar\beta_1\beta_2}\big|0\right\rangle \\
&\quad = \delta^{\alpha_1\bar\beta_1}\,\delta^{\beta_2\bar\alpha_2}\,\tilde\delta(\tilde p_1 - p_1)\,\tilde\delta(\tilde p_2 - p_2). \tag{5.40}
\end{aligned}$$

The simplicity of these Feynman rules arises due to the unconventional normalization of

the operators in Eqs. (5.37) and (5.38). This normalization has been chosen to simplify the matching of QCD amplitudes onto SCET operators, as will be seen in Sec. 5.4.

We will also make use of leptonic versions of the above currents. These are defined analogously,

$$
\begin{aligned}
J_{ij+} &= \frac{\sqrt{2}\,\varepsilon_-^\mu(n_i, n_j)}{\sqrt{\omega_i\,\omega_j}}\, \frac{\bar{\ell}_{i+}\,\gamma_\mu\ell_{j+}}{\langle n_i n_j\rangle}\,, \\
J_{ij-} &= -\frac{\sqrt{2}\,\varepsilon_+^\mu(n_i, n_j)}{\sqrt{\omega_i\,\omega_j}}\, \frac{\bar{\ell}_{i\pm}\,\gamma_\mu\ell_{j\pm}}{[n_i n_j]}\,.
\end{aligned}
\tag{5.41}
$$

Unlike the collinear quark field χ, the leptonic field ℓ does not carry color and so does not contain a strong-interaction Wilson line.

All couplings in the SM, except to the Higgs boson, preserve chirality. This limits the need for the scalar current, especially when considering only massless external quarks. In the SM the scalar current can arise through explicit couplings to the Higgs, in which case, even though we still treat the external quarks as massless, the Wilson coefficient for the scalar operator will contain the quark Yukawa coupling. This is relevant for example for $H b\bar{b}$ processes. The scalar current can also arise through off-diagonal CKM-matrix elements connecting two massless external quarks through a massive quark appearing in a loop. This can occur in multiple vector boson production, or from electroweak loop corrections, neither of which will be discussed in this chapter. When constructing an operator basis in Sec. 5.3.2, we ignore the scalar current, as it is not relevant for the examples that we will treat in this chapter. However, it should be clear that the construction of the basis in Sec. 5.3.2 can be trivially generalized to incorporate the scalar current if needed.

5.3.2 Helicity Operator Basis

Using the definitions for the gluon and quark helicity fields in Eqs. (5.26) and (5.37), we can construct operators for a given number of external partons with definite helicities and color. (As discussed at the end of the previous section, for the processes we consider in this chapter we do not require the scalar current J_S.) In the general case with CKM-matrix elements, we

232

must allow for the two quark flavors within a single current to be different. The situation is simplified in QCD processes, where one can restrict to currents carrying a single flavor label.

For an external state with n particles of definite helicities $\pm$, colors a_i, α_i, $\bar{\alpha}_i$, and flavors f, f', ..., a complete basis of operators is given by

$$
\begin{aligned}
O_{\pm\pm\cdots(\pm\cdots;\cdots\pm)}^{a_1 a_2 \cdots \bar{\alpha}_{i-1}\alpha_i \cdots \bar{\alpha}_{n-1}\alpha_n} & (\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_{i-1}, \tilde{p}_i, \ldots, \tilde{p}_{n-1}, \tilde{p}_n) \\
&= S\, \mathcal{B}_{1\pm}^{a_1} \mathcal{B}_{2\pm}^{a_2} \cdots J_{f\,i-1,i\pm}^{\bar{\alpha}_{i-1}\alpha_i} \cdots J_{f'\,n-1,n\pm}^{\bar{\alpha}_{n-1}\alpha_n}\,.
\end{aligned}
\tag{5.42}
$$

For example, $f = q$ indicates that both quark fields in the current have flavor q. When it is necessary to distinguish different flavors with the same current, for example when we consider processes involving W bosons in Sec. 5.6, we use a label $f = \bar{u}d$ such as $J_{\bar{u}d12-}$. For simplicity, we will also often suppress the dependence of the operator on the label momenta $\tilde{p}_i$. For the operator subscripts, we always put the helicity labels of the gluons first and those of the quark currents in brackets, with the labels for quark currents with different flavor labels f and f' separated by a semicolon, as in Eq. (5.42). The $\pm$ helicity labels of the individual gluon fields and quark currents can all vary independently. Operators with nonzero matching coefficients are restricted to the color-conserving subspace. We will discuss the construction of the color basis in Sec. 5.3.4.

The symmetry factor S in Eq. (5.42) is included to simplify the matching. It is given by

$$
S = \frac{1}{\prod_i n_i^+!\, n_i^-!}\,,
\tag{5.43}
$$

where $n_i^\pm$ denotes the number of fields of type $i = g, u, \bar{u}, d, \bar{d}, \ldots$ with helicity $\pm$. We also use

$$
n = \sum_i (n_i^+ + n_i^-)\,,
\tag{5.44}
$$

to denote the total number of fields in the operator. Each $\mathcal{B}_i$ counts as one field, and each J has two fields.

For each set of external particles of definite helicities, colors, and flavors, there is only one

233

independent operator, since the physical external states have been completely specified. All Feynman diagrams contributing to this specific external state will be included in the Wilson coefficient of that specific operator. For the case of pure QCD, quarks always appear in pairs of the same flavor and same chirality, and therefore can be assembled into quark currents labeled by a single flavor. In this case, to keep track of the minimal number of independent operators, we can simply order the helicity labels, and only consider operators of the form

$$O_{+\cdots(\cdots-)} = O_{\underbrace{+\cdots+}_{n_g^+}\underbrace{-\cdots-}_{n_g^-}(\underbrace{+\cdots+}_{n_q^+}\underbrace{-\cdots-}_{n_q^-})}, \tag{5.45}$$

and analogously for any additional quark currents with different quark flavors.[6]

With the operator basis constructed, for a given n-parton process we can match hard scattering processes in QCD onto the leading-power hard-scattering Lagrangian

$$\mathcal{L}_{\text{hard}} = \int \prod_{i=1}^{n} \mathrm{d}\tilde{p}_i \, C^{\alpha_1\cdots\alpha_n}_{+\cdots(\cdots-)}(\tilde{p}_1, \ldots, \tilde{p}_n) \, O^{\alpha_1\cdots\alpha_n}_{+\cdots(\cdots-)}(\tilde{p}_1, \ldots, \tilde{p}_n), \tag{5.46}$$

where a sum over all color indices is implicit. Lorentz invariance implies that the Wilson coefficients only depend on Lorentz invariant combinations of the momenta. This hard Lagrangian is used in conjunction with the collinear and soft Lagrangians that describe the dynamics of the soft and collinear modes; see for example Eq. (5.25).

We emphasize that Eq. (5.46) provides a complete basis in SCET for well-separated jets and additional nonhadronic particles at leading power. We will discuss in more detail in Sec. 5.4 the matching and regularization schemes, and demonstrate that no evanescent operators are generated for this case. At subleading power, the SCET operators would

---

[6]In the general case with off-diagonal CKM-matrix elements, there is some more freedom in the choice of the operator basis, because quarks of the same flavor do not necessarily appear in pairs. However, it is still true that only a single operator is needed for a specific external state. For example, for external quarks $u_-, \bar{d}_+, \bar{s}_+, c_-$, one could either use the operators $J_{us-}J_{cd-}$, or the operators $J_{cs-}J_{ud-}$ (where the color structures have been suppressed). Since different helicity combinations are possible, a single flavor assignment does not suffice to construct a complete helicity basis, and one must sum over a basis of flavor assignments. As an example explicitly demonstrating this, we will consider the case of $pp \to W+$ jets in Sec. 5.6.

involve additional derivative operators, soft fields, or multiple SCET building blocks from the same collinear sector.

5.3.3 Example with a Z-Boson Exchange

It is important to note that all kinematic dependence of the hard process, for example, its angular distributions, is encoded in the Wilson coefficients. Since the Wilson coefficients can (in principle) carry an arbitrary kinematic dependence, our choice of helicity basis imposes no restriction on the possible structure or mediating particles of the hard interaction. For example, the spin of an intermediate particle may modify the angular distribution of the decay products, and hence the Wilson coefficients, but this can always be described by the same basis of helicity operators.

As a simple example to demonstrate this point we consider $e^+ e^- \to e^+ e^-$ at tree level. This process can proceed through either an off-shell γ or Z boson. Because the SM couplings to both of these particles preserve chirality, a basis of operators for this process is given by

$$
O_{(++)} = \frac{1}{4} J_{e\,12+}\, J_{e\,34+}\,,
$$

$$
O_{(+-)} = J_{e\,12+}\, J_{e\,34-}\,,
$$

$$
O_{(--)} = \frac{1}{4} J_{e\,12-}\, J_{e\,34-}\,, \tag{5.47}
$$

where the leptonic current is defined in Eq. (5.41). The fact that this is a complete basis relies only on the fact that the couplings preserve chirality, and is independent of e.g. the possible number of polarizations of the mediating Z or γ.

We now consider the calculation of the Wilson coefficients for the matching to these operators (the matching procedure is discussed in detail in Sec. 5.4). At tree level, the Wilson coefficients are easily calculated, giving

$$
C_{(++)} = -e^2 \big[1 + v_R^e v_R^e P_Z(s_{12}) \big] \frac{2[13]\langle 24\rangle}{s_{12}} + (1 \leftrightarrow 3)\,,
$$

$$
C_{(+-)} = -e^2 \big[1 + v_R^e v_L^e P_Z(s_{12}) \big] \frac{2[14]\langle 23\rangle}{s_{12}}\,,
$$

$$C_{(--)} = -e^2 \big[1 + v_L^e v_L^e P_Z(s_{12}) \big] \frac{2[24]\langle 13 \rangle}{s_{12}} + (1 \leftrightarrow 3) \,. \tag{5.48}$$

Here $s_{12} = (p_1 + p_2)^2$, P_Z is the ratio of the Z and photon propagators,

$$P_Z(s) = \frac{s}{s - m_Z^2 + i\Gamma_Z m_Z} \,, \tag{5.49}$$

and the couplings $v_{L,R}$ to the Z boson are

$$v_L^e = \frac{1 - 2\sin^2\theta_W}{\sin(2\theta_W)} \,, \quad v_R^e = -\frac{2\sin^2\theta_W}{\sin(2\theta_W)} \,. \tag{5.50}$$

Note that the presence of the spinor factors in Eq. (5.48) occur due to our normalization conventions for the currents.

Now, consider calculating the scattering amplitude in the effective theory, for example for the case when both electrons have positive helicity. The matrix element in the effective theory gives

$$\Big\langle e_+^-(p_1) \, e_-^+(p_2) \, e_+^-(p_3) \, e_-^+(p_4) \big| i\mathcal{L}_{\text{hard}} \big| 0 \Big\rangle$$

$$= i\Big\langle e_+^-(p_1) \, e_-^+(p_2) \, e_+^-(p_3) \, e_-^+(p_4) \Big| \int \prod_{i=1}^{n} d\tilde{p}_i \, C_{++} O_{++} \big| 0 \Big\rangle$$

$$= -ie^2 \big[1 + v_R^e v_R^e P_Z(s_{12}) \big] \frac{2[13]\langle 24 \rangle}{s_{12}} + (1 \leftrightarrow 3) \,, \tag{5.51}$$

using the Feynman rules of Eq. (5.40). The effective theory therefore reproduces the full theory scattering amplitude. The same is true of the other helicity configurations, so the familiar angular distributions for $e^+e^- \to e^+e^-$, as well as the different couplings of the Z to left- and right-handed particles, are entirely encoded in the Wilson coefficients.

5.3.4 Color Basis

In addition to working with a basis of operators with definite helicity, we can also choose a color basis that facilitates the matching. When constructing a basis of operators in SCET, we

are free to choose an arbitrary color basis. With respect to color, we can think of Eq. (5.46) as having a separate Wilson coefficient for each color configuration. For specific processes the color structure of the Wilson coefficients can be further decomposed as

$$C^{a_1 \cdots \alpha_n}_{+\cdots(\cdots-)} = \sum_k C^k_{+\cdots(\cdots-)} T^{a_1 \cdots \alpha_n}_k \equiv \bar{T}^{a_1 \cdots \alpha_n} \vec{C}_{+\cdots(\cdots-)}. \tag{5.52}$$

Here, $\bar{T}^{a_1 \cdots \alpha_n}$ is a row vector whose entries $T^{a_1 \cdots \alpha_n}_k$ are suitable color structures that together provide a complete basis for all allowed color structures, but which do not necessarily all have to be independent. In other words, the elements of $\bar{T}^{a_1 \cdots \alpha_n}$ span the color-conserving subspace of the full color space spanned by $\{a_1 \cdots \alpha_n\}$, and $\vec{C}$ is a vector in this subspace. Throughout this chapter we will refer to the elements of $\bar{T}^{a_1 \cdots \alpha_n}$ as a color basis, although they will generically be overcomplete, since this allows for simpler choices of color structures. As discussed below, due to the overcompleteness of the bases, some care will be required for their consistent usage.

Using Eq. (5.52), we can rewrite Eq. (5.46) as

$$\mathcal{L}_{\text{hard}} = \int \prod_{i=1}^n \mathrm{d}\tilde{p}_i \, \vec{O}^\dagger_{+\cdots(\cdots-)}(\tilde{p}_1, \ldots, \tilde{p}_n) \vec{C}_{+\cdots(\cdots-)}(\tilde{p}_1, \ldots, \tilde{p}_n), \tag{5.53}$$

where $\vec{O}^\dagger$ is a conjugate vector defined by

$$\vec{O}^\dagger_{+\cdots(\cdots-)} = O^{a_1 \cdots \alpha_n}_{+\cdots(\cdots-)} \bar{T}^{a_1 \cdots \alpha_n}. \tag{5.54}$$

While the form $C^{a_1 \cdots \alpha_n}_{+\cdots(\cdots-)} O^{a_1 \cdots \alpha_n}_{+\cdots(\cdots-)}$ in Eq. (5.46) is more convenient to discuss the matching and the symmetry properties of operators and Wilson coefficients, the alternative form in Eq. (5.53) is more convenient to discuss the mixing of the color structures under renormalization.

For low multiplicities of colored particles it can be convenient to use orthogonal color bases, e.g., the singlet-octet basis for $q\bar{q}q'\bar{q}'$ is orthogonal. However, using orthogonal bases becomes increasingly difficult for higher multiplicity processes, and the color bases used for

many fixed-order calculations are not orthogonal. (See e.g. Refs. [333, 334] for a discussion of the use of orthogonal bases for $SU(N)$.) The use of a nonorthogonal color basis implies that when written in component form in a particular basis, the conjugate $\vec{C}^\dagger$ of the vector $\vec{C}$ is not just given by the naive complex conjugate transpose of the components of the vector. Instead, we have

$$\vec{C}^\dagger = [C^{a_1 \cdots a_n}]^* \, \bar{T}^{a_1 \cdots a_n} = \vec{C}^{*T} \, \widehat{T} \,, \tag{5.55}$$

where

$$\widehat{T} = \sum_{a_1, \ldots, a_n} (\bar{T}^{a_1 \cdots a_n})^\dagger \bar{T}^{a_1 \cdots a_n} \tag{5.56}$$

is the matrix of color sums for the chosen basis. If the basis is orthogonal (orthonormal), then $\widehat{T}$ is a diagonal matrix (identity matrix). Note that Eq. (5.56) implies that by definition $\widehat{T}^{*T} = \widehat{T}$.

Similar to Eq. (5.55), for an abstract matrix $\widehat{X}$ in color space, the components of its Hermitian conjugate $\widehat{X}^\dagger$ when written in a particular basis are given in terms of the components of $\widehat{X}$ as

$$\widehat{X}^\dagger = \widehat{T}^{-1} \, \widehat{X}^{*T} \, \widehat{T} \,. \tag{5.57}$$

A proper treatment of the nonorthogonality of the color basis is also important in the factorization theorem of Eq. (5.2). Here, the color indices of the Wilson coefficients are contracted with the soft function as

$$[C^{a_1 \cdots a_n}]^* \, S_\kappa^{a_1 \cdots a_n b_1 \cdots \beta_n} C^{b_1 \cdots \beta_n} = \vec{C}^\dagger \, \widehat{S}_\kappa \vec{C}$$
$$= \vec{C}^{*T} \widehat{T} \, \widehat{S}_\kappa \vec{C} \,. \tag{5.58}$$

At tree level, the soft function is simply the color-space identity

$$\widehat{S}_\kappa = \mathbf{1} \,, \tag{5.59}$$

238

which follows from its color basis independent definition in terms of Wilson lines [see e.g. Ref. [301] or Eq. (5.267)]. Here we have suppressed the dependence of $\widehat{S}$ on soft momenta. The action of the identity on an element of the color space is defined by

$$(1\bar{T})^{\cdots a_i \cdots \alpha_j \cdots} = \bar{T}^{\cdots a_i \cdots \alpha_j \cdots} \,, \tag{5.60}$$

and its matrix representation in any color basis is given by $\mathbf{1} = \mathrm{diag}(1, 1, \cdots, 1)$. In the literature, see e.g. Refs. [335, 296, 290, 297, 323], often a different convention is used, where the $\widehat{T}$ matrix is absorbed into the definition of the soft function. In this convention, the soft function becomes explicitly basis dependent and is not the same as the basis-independent color-space identity. One should be careful to not identify the two.

As an example to demonstrate our notation for the color basis, consider the process $gg q\bar{q}$. A convenient choice for a complete basis of color structures is

$$\bar{T}^{ab\alpha\bar{\beta}} = \left((T^a T^b)_{\alpha\bar{\beta}} \,, \, (T^b T^a)_{\alpha\bar{\beta}} \,, \, \mathrm{tr}[T^a T^b]\, \delta_{\alpha\bar{\beta}} \right)$$

$$\equiv \begin{pmatrix} (T^a T^b)_{\alpha\bar{\beta}} \\ (T^b T^a)_{\alpha\bar{\beta}} \\ \mathrm{tr}[T^a T^b]\, \delta_{\alpha\bar{\beta}} \end{pmatrix}^T \,. \tag{5.61}$$

For cases with many color structures we will write $\bar{T}$ as the transpose of a column vector as above. The transpose in this case only refers to the vector itself, not to the individual color structures. The color-sum matrix for this particular basis is

$$\widehat{T}_{gg q\bar{q}} = (\bar{T}^{ab\alpha\bar{\beta}})^\dagger \bar{T}^{ab\alpha\bar{\beta}}$$

$$= \frac{C_F N}{2} \begin{pmatrix} 2C_F & 2C_F - C_A & 2T_F \\ 2C_F - C_A & 2C_F & 2T_F \\ 2T_F & 2T_F & 2T_F N \end{pmatrix} \,. \tag{5.62}$$

Our conventions for color factors are given in Appendix B.1.2. Explicit expressions for $\widehat{T}$ for

the bases used in this chapter are given in Appendix B.6 for up to five partons.

Depending on the application, different choices of color basis can be used. For example, in fixed-order QCD calculations, color ordering [226, 227, 228, 229] is often used to organize color information and simplify the singularity structure of amplitudes, while the color flow basis [336] is often used to interface with Monte Carlo generators. For a brief review of the color decomposition of QCD amplitudes, see Appendix B.1.3. Choosing a corresponding color basis in SCET has the advantage that the Wilson coefficients are given directly by the finite parts of the color-stripped helicity amplitudes, as defined in Eq. (5.96), which can be efficiently calculated using unitarity methods. In this chapter we will use color bases corresponding to the color decompositions of the QCD amplitudes when giving explicit results for the matching coefficients, although we emphasize that an arbitrary basis can be chosen depending on the application.

Finally, note that the color structures appearing in the decomposition of a QCD amplitude up to a given loop order may not form a complete basis. The color basis in SCET must be complete even if the matching coefficients of some color structures are zero to a given loop order, since all structures can in principle mix under renormalization group evolution, as will be discussed in Sec. 5.8. In this case, we always choose a complete basis in SCET such that the color structures appearing in the amplitudes to some fixed order are contained as a subset.

5.3.5 Parity and Charge Conjugation

Under charge conjugation, the fields transform as

$$C \, \mathcal{B}_{i\pm}^a \, T_{\alpha\bar{\beta}}^a \, C = -\mathcal{B}_{i\pm}^a T_{\beta\bar{\alpha}}^a \, ,$$
$$C \, J_{ij\pm}^{\bar{\alpha}\beta} \, C = -J_{ji\mp}^{\bar{\beta}\alpha} \, . \tag{5.63}$$

The minus sign on the right-hand side of the second equation comes from anticommutation of the fermion fields.

Under parity, the fields transform as

$$P \, \mathcal{B}_{i\pm}^{a}(\tilde{p}_i, x) \, P = e^{\pm 2i\phi_{n_i}} \mathcal{B}_{i\mp}^{a}(\tilde{p}_i^{P}, x^{P}) \,,$$

$$P \, J_{ij\pm}^{\bar{\alpha}\beta}(\tilde{p}_i, \tilde{p}_j, x) \, P = e^{\pm i(\phi_{n_i} - \phi_{n_j})} J_{ij\mp}^{\bar{\alpha}\beta}(\tilde{p}_i^{P}, \tilde{p}_j^{P}, x^{P}) \,, \tag{5.64}$$

where we have made the dependence on $\tilde{p}_i$ and x explicit, and the parity-transformed vectors are $\tilde{p}_i^{P} = \omega_i \, \bar{n}_i/2$, $x_\mu^{P} = x^\mu$. The ϕ_{n_i} are real phases, whose exact definition is given in Appendix B.1.1. The phases appearing in the parity transformation of the helicity operators exactly cancel the phases appearing in the corresponding helicity amplitude under a parity transformation. This overall phase is determined by the little group scaling (see Appendix B.1.1 for a brief review).

Using the transformations of the helicity fields under parity and charge conjugation in Eqs. (5.63) and (5.64), it is straightforward to determine how these discrete symmetries act on the helicity operators. Parity and charge conjugation invariance of QCD implies that the effective Lagrangian in Eq. (5.46) must also be invariant. (For amplitudes involving electroweak interactions, parity and charge conjugation invariance are explicitly violated. This is treated by extracting parity and charge violating couplings from the operators and amplitudes. See Sec. 5.6 for a discussion.) This then allows one to derive corresponding relations for the Wilson coefficients.

To illustrate this with a nontrivial example we consider the $gg q\bar{q}$ process. The operators transform under charge conjugation as

$$\begin{aligned}
C \, O_{\lambda_1 \lambda_2(\pm)}^{ab\,\bar{\alpha}\beta}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) \, \bar{T}^{ab\alpha\bar{\beta}} \, C \\
= C \, S \, \mathcal{B}_{1\lambda_1}^{a} \mathcal{B}_{2\lambda_2}^{b} J_{34\pm}^{\bar{\alpha}\beta} \, \bar{T}^{ab\alpha\bar{\beta}} \, C \\
= -O_{\lambda_1 \lambda_2(\mp)}^{ba\,\bar{\alpha}\beta}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_4, \tilde{p}_3) \, \bar{T}^{ab\alpha\bar{\beta}} \,, \tag{5.65}
\end{aligned}$$

where $\lambda_{1,2}$ denote the gluon helicities, and $\bar{T}^{ab\alpha\beta}$ is as given in Eq. (5.61). From the invariance

of Eq. (5.46) we can infer that the Wilson coefficients must satisfy

$$C^{ab\,\alpha\bar\beta}_{\lambda_1\lambda_2(\pm)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4) = -C^{ba\,\alpha\bar\beta}_{\lambda_1\lambda_2(\mp)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_4,\tilde{p}_3)\,. \qquad (5.66)$$

In the color basis of Eq. (5.61), we can write this as

$$\vec{C}_{\lambda_1\lambda_2(\pm)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4) = \widehat{V}\vec{C}_{\lambda_1\lambda_2(\mp)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_4,\tilde{p}_3)\,,$$

$$\text{with}\qquad \widehat{V} = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \qquad (5.67)$$

Now consider the behavior under parity. For concreteness we consider the case of positive helicity gluons. The operators transform as

$$\mathrm{P}\,O^{ab\,\bar\alpha\beta}_{++(\pm)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4)\,\mathrm{P} = \mathrm{P}\,\frac{1}{2}\mathcal{B}^a_{1+}\mathcal{B}^b_{2+}J^{\bar\alpha\beta}_{34\pm}\,\mathrm{P} \qquad (5.68)$$

$$= e^{\mathrm{i}(2\phi_{n_1}+2\phi_{n_2}\pm(\phi_{n_3}-\phi_{n_4}))}O^{ab\,\bar\alpha\beta}_{--(\mp)}(\tilde{p}^P_1,\tilde{p}^P_2;\tilde{p}^P_3,\tilde{p}^P_4)\,.$$

The invariance of Eq. (5.46) under parity then implies that the Wilson coefficients satisfy

$$\vec{C}_{++(\pm)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4)$$

$$= \vec{C}_{--(\mp)}(\tilde{p}^P_1,\tilde{p}^P_2;\tilde{p}^P_3,\tilde{p}^P_4)e^{-\mathrm{i}(2\phi_{n_1}+2\phi_{n_2}\pm(\phi_{n_3}-\phi_{n_4}))}$$

$$= \vec{C}_{--(\mp)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4)\Big|_{\langle..\rangle\leftrightarrow[..]}\,. \qquad (5.69)$$

Here we have introduced the notation $\langle..\rangle \leftrightarrow [..]$ to indicate that all angle and square spinors have been switched in the Wilson coefficient. The fact that the phase appearing in the parity transformation of the operator exactly matches the phase arising from evaluating the Wilson coefficient with parity related momenta is guaranteed by little group scaling, and will therefore occur generically. See Eqs. (B.24) and (B.25) and the surrounding discussion for a review.

Below we will use charge conjugation to reduce the number of Wilson coefficients for which we have to carry out the matching explicitly. We will use parity only when it helps to avoid substantial repetitions in the matching.

5.3.6 Crossing Symmetry

Our basis is automatically crossing symmetric, since the gluon fields $\mathcal{B}_{i\pm}$ can absorb or emit a gluon and the quark current $J_{ij\pm}$ can destroy or produce a quark-antiquark pair, or destroy and create a quark or antiquark. We will first illustrate how to use crossing symmetry in an example and then describe how to technically have crossing symmetric Wilson coefficients.

We will again consider the process $gg q\bar{q}$ as an example. Due to our outgoing conventions, the default Wilson coefficient is for the unphysical processes with all outgoing particles:

$$0 \rightarrow g_+^a(p_1) g_-^b(p_2) q_+^\alpha(p_3) \bar{q}_-^{\bar\beta}(p_4) : C_{+-(+)}^{ab\alpha\bar\beta}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) , \qquad (5.70)$$

where we picked one specific helicity configuration for definiteness. Crossing a particle from the final state to the initial state flips its helicity, changes the sign of its momentum, and changes it to its antiparticle. In addition we get a minus sign for each crossed fermion, though in practice these can be ignored as they do not modify the cross section. This allows one to obtain the Wilson coefficient for any crossing. For example, for the following possible crossings, the Wilson coefficients are given by

$$
\begin{aligned}
g_+^a(p_1) g_-^b(p_2) &\rightarrow q_+^\alpha(p_3) \bar{q}_-^{\bar\beta}(p_4) : C_{+-(+)}^{ba\alpha\bar\beta}(-\tilde{p}_2, -\tilde{p}_1; \tilde{p}_3, \tilde{p}_4) , \\
g_+^a(p_1) q_+^{\bar\alpha}(p_2) &\rightarrow g_+^b(p_3) q_+^\beta(p_4) : -C_{+-(+)}^{ba\beta\bar\alpha}(\tilde{p}_3, -\tilde{p}_1; \tilde{p}_4, -\tilde{p}_2) , \\
g_+^a(p_1) \bar{q}_-^\alpha(p_2) &\rightarrow g_+^b(p_3) \bar{q}_-^{\bar\beta}(p_4) : -C_{+-(+)}^{ba\alpha\bar\beta}(\tilde{p}_3, -\tilde{p}_1; -\tilde{p}_2, \tilde{p}_4) , \\
q_+^{\bar\alpha}(p_1) \bar{q}_-^\beta(p_2) &\rightarrow g_+^a(p_3) g_-^b(p_4) : C_{+-(+)}^{ab\beta\bar\alpha}(\tilde{p}_3, \tilde{p}_4; -\tilde{p}_2, -\tilde{p}_1) .
\end{aligned} \qquad (5.71)
$$

Since the signs of momenta change when crossing particles between the final and initial state, care is required in taking the proper branch cuts to maintain crossing symmetry for

the Wilson coefficients. In terms of the Lorentz invariants

$$s_{ij} = (p_i + p_j)^2 \tag{5.72}$$

this amounts to the choice of branch cut defined by $s_{ij} \to s_{ij} + i0$. In particular, we write all logarithms as

$$L_{ij} \equiv \ln\left(-\frac{s_{ij}}{\mu^2} - i0\right) = \ln\left(\frac{s_{ij}}{\mu^2}\right) - i\pi\theta(s_{ij}). \tag{5.73}$$

For spinors, crossing symmetry is obtained by defining the conjugate spinors $\langle p\pm|$ as was done in Eq. (5.5), resulting in the following relation

$$\langle p\pm| = \text{sgn}(p^0)\,\overline{|p\pm\rangle}. \tag{5.74}$$

The additional minus sign for negative p^0 is included to use the same branch (of the square root inside the spinors) for both spinors and conjugate spinors, i.e., for $p^0 > 0$ we have

$$|(-p)\pm\rangle = i\,|p\pm\rangle\,,$$
$$\langle(-p)\pm| = -(-i)\,\langle p\pm| = i\,\langle p\pm|\,. \tag{5.75}$$

In this way all spinor identities are automatically valid for both positive and negative momenta, which makes it easy to use crossing symmetry.

5.3.7 Hard Function

In the factorized expression for the cross section given in Eq. (5.2), the dependence on the underlying hard Born process appears through the hard function $\widehat{H}_\kappa$. In terms of the Wilson coefficients of the operator basis in the effective theory, the hard function for a particular partonic channel κ is given by

$$\widehat{H}_\kappa(\{\tilde{p}_i\}) = \sum_{\{\lambda_i\}} \vec{C}_{\lambda_1\cdots(\cdot\lambda_n)}(\{\tilde{p}_i\})\,\vec{C}^\dagger_{\lambda_1\cdots(\cdot\lambda_n)}(\{\tilde{p}_i\})\,, \tag{5.76}$$

where $\{\tilde{p}_i\} \equiv \{\tilde{p}_1, \tilde{p}_2, \ldots\}$. For unpolarized experiments we simply sum over all helicity operators, so $\widehat{H}_\kappa(\{\tilde{p}_i\})$ with its sum over helicities in Eq. (5.76) appears as a multiplicative factor. It is important to note that the color indices of the Wilson coefficients are not contracted with each other, rather they are contracted with the color indices of the soft function through the trace seen in Eq. (5.2).

As an explicit example to demonstrate the treatment of both color and helicity indices, we consider the contribution of the $gg q\bar{q}$ partonic channel to the $pp \to 2$ jets process. In this case, the Wilson coefficients are given by $\vec{C}_{\lambda_1 \lambda_2 (\lambda_3)}$, where λ_1, λ_2 denote the helicities of the gluons, λ_3 denotes the helicity of the quark current, and recall that the vector denotes the possible color structures, which were given explicitly for this case in Eq. (5.61). The hard function for this partonic channel is then given by

$$
\begin{aligned}
\widehat{H}_{gg q\bar{q}}(\{\tilde{p}_i\}) = &\sum_{\{\lambda_i\}} \vec{C}_{\lambda_1 \lambda_2 (\lambda_3)}(\{\tilde{p}_i\}) \, \vec{C}^{\dagger}_{\lambda_1 \lambda_2 (\lambda_3)}(\{\tilde{p}_i\}) \\
= &\vec{C}_{++(+)} \vec{C}^{\dagger}_{++(+)} + \vec{C}_{+-(+)} \vec{C}^{\dagger}_{+-(+)} + \\
&\vec{C}_{-+(+)} \vec{C}^{\dagger}_{-+(+)} + \vec{C}_{--(+)} \vec{C}^{\dagger}_{--(+)} + \\
&\vec{C}_{++(-)} \vec{C}^{\dagger}_{++(-)} + \vec{C}_{+-(-)} \vec{C}^{\dagger}_{+-(-)} + \\
&\vec{C}_{-+(-)} \vec{C}^{\dagger}_{-+(-)} + \vec{C}_{--(-)} \vec{C}^{\dagger}_{--(-)} .
\end{aligned}
\tag{5.77}
$$

Here, explicit expressions are only needed for $\vec{C}_{++(+)}, \vec{C}_{+-(+)}$ and $\vec{C}_{--(+)}$. One can obtain $\vec{C}_{-+(+)}$ using Bose symmetry simply by interchanging the gluons,

$$
\vec{C}^{ab\alpha\bar{\beta}}_{-+(+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) = \vec{C}^{ba\alpha\bar{\beta}}_{+-(+)}(\tilde{p}_2, \tilde{p}_1; \tilde{p}_3, \tilde{p}_4) ,
\tag{5.78}
$$

or equivalently,

$$
\vec{C}_{-+(+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) = \widehat{V} \vec{C}_{+-(+)}(\tilde{p}_2, \tilde{p}_1; \tilde{p}_3, \tilde{p}_4) ,
$$

$$\text{with} \qquad \widehat{V} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{5.79}$$

As explained in Sec. 5.7.1, the remaining $\vec{C}_{\lambda_1 \lambda_2(-)}$ can be obtained from the expressions for the other Wilson coefficients by charge conjugation.

In Eq. (5.77), the Wilson coefficients are vectors in the color basis of Eq. (5.61) and thus the hard function is a matrix in this basis. As discussed in Sec. 5.3.4, the tree-level soft function is the color-space identity, i.e.,

$$S_{gg q\bar{q}}^{(0)\, b_1 b_2 \beta_1 \bar{\beta}_2 \, a_1 a_2 \alpha_1 \bar{\alpha}_2} = \delta^{b_1 a_1} \delta^{b_2 a_2} \delta_{\beta_1 \alpha_1} \delta_{\bar{\beta}_2 \bar{\alpha}_2} \equiv \mathbf{1} . \tag{5.80}$$

With the color trace in Eq. (5.2) this amounts to contracting the color indices of the Wilson coefficients. In the color basis of Eq. (5.61), this simply becomes

$$\widehat{S}_{gg q\bar{q}}^{(0)} = \mathbf{1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{5.81}$$

The tree-level soft function also has dependence on momenta depending on the measurement being made, which are not shown here.

To demonstrate a complete calculation of the cross section using the factorization theorem of Eq. (5.2) together with the hard functions computed using the helicity operator formalism, it is instructive to see how the leading-order cross section is reproduced from Eq. (5.2). We consider the simple case of $H + 0$ jets in the $m_t \to \infty$ limit. For this channel, there is a unique color structure $\delta_{a_1 a_2}$, and using the results of Sec. 5.5.1 and Appendix B.2.1, the lowest order Wilson coefficients are given by

$$\vec{C}_{++}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3) = \delta_{a_1 a_2} \frac{\alpha_s}{3\pi v} \frac{s_{12}}{2} \frac{[12]}{\langle 12 \rangle}, \tag{5.82}$$

$$\vec{C}_{--}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3) = \delta_{a_1 a_2} \frac{\alpha_s}{3\pi v} \frac{s_{12}}{2} \frac{\langle 12 \rangle}{[12]}, \tag{5.83}$$

$$\vec{C}_{+-}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3) = \vec{C}_{-+}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3) = 0, \tag{5.84}$$

where $v = \left(\sqrt{2} G_F\right)^{-1/2} = 246 \text{GeV}$. Note that these are simply the helicity amplitudes for the process, as will be shown more generally in Sec. 5.4. Analytically continuing to physical momenta, squaring, and summing over helicities, the tree-level hard function is given by

$$\begin{aligned} H_{ggH}^{(0)\, a_1 a_2 b_1 b_2}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3) &= \left| \frac{\alpha_s}{3\pi v} \frac{s_{12}}{2} \right|^2 2\, \delta_{a_1 a_2} \delta_{b_1 b_2} \\ &= \frac{\alpha_s^2 s_{12}^2}{18\pi^2 v^2}\, \delta_{a_1 a_2} \delta_{b_1 b_2}\,. \end{aligned} \tag{5.85}$$

Note that only 2 of the 4 helicity configurations contribute, hence the factor of 2.

The tree-level gluon beam functions are given by the gluon PDFs. Since there are no jets in the final state, there are no jet functions. The tree-level soft function is the identity in color space[7]

$$S_{gg}^{(0)\, b_1 b_2 a_1 a_2} = \delta^{b_1 a_1} \delta^{b_2 a_2}\,. \tag{5.87}$$

The leading-order cross section is then given by

$$\begin{aligned} \sigma &= \frac{1}{2 E_{\text{cm}}^2} \frac{1}{[2(N^2-1)]^2} \int \frac{\mathrm{d}x_1}{x_1} \frac{\mathrm{d}x_2}{x_2}\, f_g(x_1) f_g(x_2) \\ &\quad \times \int \frac{\mathrm{d}^4 p_3}{(2\pi)^3}\, \theta(p_3^0)\, \delta(p_3^2 - m_H^2) \\ &\quad \times (2\pi)^4 \delta^4\!\left(x_1 E_{\text{cm}} \frac{n_1}{2} + x_2 E_{\text{cm}} \frac{n_2}{2} - p_3 \right) \\ &\quad \times H_{ggH}^{(0)\, a_1 a_2 b_1 b_2}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3)\, S_{gg}^{(0)\, b_1 b_2 a_1 a_2} \\ &= \frac{\alpha_s^2 m_H^2}{576\pi v^2 E_{\text{cm}}^2} \int \mathrm{d}Y\, f_g\!\left(\frac{m_H}{E_{\text{cm}}} e^Y \right) f_g\!\left(\frac{m_H}{E_{\text{cm}}} e^{-Y} \right). \end{aligned} \tag{5.88}$$

---

[7]Since there is only one color structure, the tree-level soft function is normally defined as

$$S_{gg}^{(0)} = \frac{1}{N^2 - 1}\, \delta_{a_1 a_2} \delta_{b_1 b_2}\, \delta^{b_1 a_1} \delta^{b_2 a_2} = 1\,. \tag{5.86}$$

Here we do not absorb numerical prefactors into our soft functions, because this is not useful for processes with more final-state partons.

The $1/(2E_{\rm cm}^2)$ factor is the flux factor and for each of the incoming gluons we get a $1/[2(N^2 - 1)]$ from averaging over its spin and color. This is followed by integrals over the gluon PDFs, f_g, and the Higgs phase space, where we have restricted to the production of an on-shell Higgs. The final expression in Eq. (5.88) agrees with the standard result, where the first factor is the Born cross section.

We now briefly discuss our choice of normalization. The currents in Eq. (5.37) were normalized such that the Wilson coefficients are simply given by the finite part of the QCD helicity amplitudes (see Eq. (5.96) and Sec. 5.4). This is distinct from the normalization typically used for SCET operators, e.g. $\bar{\chi}_i\gamma^\mu\chi_j$, which is chosen to facilitate the matching to QCD operators. We now show that the extra factors in Eq. (5.37) arrange themselves to produce the standard normalization for the jet function (or beam function). Starting from the current and its conjugate,

$$J_{ij\pm}^{\bar\alpha\beta}(J_{ij\pm}^{\bar\gamma\delta})^\dagger \tag{5.89}$$

$$= \frac{\sqrt{2}\varepsilon_\mp^\mu(n_i,n_j)}{\sqrt{\omega_i\omega_j}} \frac{\bar\chi_{i\pm}^{\bar\alpha}\gamma_\mu\chi_{j\pm}^\beta}{\langle n_i\mp |n_j\pm\rangle} \frac{\sqrt{2}\varepsilon_\pm^\nu(n_i,n_j)}{(\sqrt{\omega_i\omega_j})^*} \frac{\bar\chi_{j\pm}^{\bar\delta}\gamma_\nu\chi_{i\pm}^\gamma}{\langle n_j\pm |n_i\mp\rangle}$$

$$= 4\frac{\delta^{\gamma\bar\alpha}}{N}\frac{\delta^{\beta\bar\delta}}{N}\frac{\varepsilon_\pm^\nu(n_i,n_j)\,\varepsilon_\mp^\mu(n_i,n_j)}{2n_i{\cdot}n_j\,|\omega_i\omega_j|}\,{\rm tr}\Big[\gamma_\nu\frac{\slashed{n}_i}{4}\gamma_\mu\frac{\slashed{n}_j}{4}\Big]$$

$$\times\Big(\bar\chi_{i\pm}\frac{\slashed{n}_i}{2}\chi_{i\pm}\Big)\Big(\bar\chi_{j\pm}\frac{\slashed{n}_j}{2}\chi_{j\pm}\Big) + \dots$$

$$= 2\,\delta^{\gamma\bar\alpha}\delta^{\beta\bar\delta}\Big(\frac{1}{2N}\frac{1}{|\omega_i|}\bar\chi_{i\pm}\frac{\slashed{n}_i}{2}\chi_{i\pm}\Big)\Big(\frac{1}{2N}\frac{1}{|\omega_j|}\bar\chi_{j\pm}\frac{\slashed{n}_j}{2}\chi_{j\pm}\Big)\,,$$

where we have rearranged the expression in a factorized form using the SCET Fierz formula in spin

$$1\otimes 1 = \frac{1}{2}\Big[\frac{\slashed{n}_i}{2}\otimes\frac{\slashed{n}_i}{2} - \frac{\slashed{n}_i\gamma_5}{2}\otimes\frac{\slashed{n}_i\gamma_5}{2} - \frac{\slashed{n}_i\gamma_\perp^\mu}{2}\otimes\frac{\slashed{n}_i\gamma_{\perp\mu}}{2}\Big], \tag{5.90}$$

which applies for the SCET projected spinors. In the last line of Eq. (5.89) we have dropped the color nonsinglet terms and terms which vanish when averaging over helicities, which are indicated by ellipses. The delta functions in color space highlight that the jet function does

not modify the color structure. The factor $1/\omega_{i,j}$, which arises from the normalization of the helicity currents, is part of the standard definition of the jet function and ensures that this operator has the correct mass dimension.

5.4 Matching and Scheme Dependence

In this section, we discuss the matching of QCD onto the SCET helicity operator basis introduced in the previous section. We start with a discussion of the matching for generic helicity operators in Sec. 5.4.1. In Sec. 5.4.2 we discuss in detail the subject of renormalization schemes, and the issue of converting between regularization/renormalization schemes commonly used in spinor-helicity calculations, and those used in SCET. We also demonstrate that evanescent operators are not generated in our basis.

5.4.1 Generic Matching

In this chapter, we work to leading order in the power counting, which means we only require operators that contain exactly one field per collinear sector. That is, different n_i in Eq. (5.42) are implicitly restricted to belong to different equivalence classes, $\{n_i\} \neq \{n_j\}$ for $i \neq j$. Operators with more than one field per collinear direction are power-suppressed compared to the respective leading-order operators that have the same set of collinear directions and the minimal number of fields.

At leading order, the Wilson coefficients can thus be determined by computing matrix elements of Eq. (5.46), with all external particles assigned well-separated momenta, so that they belong to separate collinear sectors. The only helicity operator that contributes in this case is the one that matches the set of external helicities, picking out the corresponding Wilson coefficient. Since we only have one external particle per collinear sector, we can simply choose $n_i = p_i/p_i^0$ in the matching calculation to represent the equivalence class $\{n_i\}$.

To compute the matrix element of $\mathcal{L}_{\text{hard}}$, we first note that the helicity operators are symmetric (modulo minus signs from fermion anticommutation) under simultaneously inter-

249

changing the label momenta and indices of identical fields, and the same is thus also true for their Wilson coefficients. For example, at tree level

$$
\begin{aligned}
\langle g_+^{a_1}(p_1) \, g_+^{a_2}(p_2) | O_{++}^{b_1 b_2} | 0 \rangle^{\text{tree}} \\
= \frac{1}{2} \big[\delta^{a_1 b_1} \, \delta^{a_2 b_2} \tilde{\delta}(\tilde{p}_1 - p_1) \, \tilde{\delta}(\tilde{p}_2 - p_2) \\
+ \delta^{a_1 b_2} \, \delta^{a_2 b_1} \tilde{\delta}(\tilde{p}_1 - p_2) \, \tilde{\delta}(\tilde{p}_2 - p_1) \big]
\end{aligned}
\tag{5.91}
$$

so the tree-level matrix element of $\mathcal{L}_{\text{hard}}$ gives

$$
\langle g_+^{a_1}(p_1) \, g_+^{a_2}(p_2) | \mathcal{L}_{\text{hard}} | 0 \rangle^{\text{tree}}
\tag{5.92}
$$
$$
= \frac{1}{2} \big[C_{++}^{a_1 a_2}(\tilde{p}_1, \tilde{p}_2) + C_{++}^{a_2 a_1}(\tilde{p}_2, \tilde{p}_1) \big] = C_{++}^{a_1 a_2}(\tilde{p}_1, \tilde{p}_2) \, .
$$

By choosing $n_i = p_i / p_i^0$, the label momenta $\tilde{p}_i$ on the right-hand side simply become $\tilde{p}_i \equiv \bar{n} \cdot p_i \, n_i / 2 = p_i$.

Taking into account the symmetry factor in Eq. (5.43), one can easily see that this result generalizes to more than two gluons or quark currents with the same helicity. In the case of identical fermions, the various terms in the operator matrix element have relative minus signs due to fermion anticommutation which precisely match the (anti)symmetry properties of the Wilson coefficients. Hence, the tree-level matrix element of $\mathcal{L}_{\text{hard}}$ is equal to the Wilson coefficient that corresponds to the configuration of external particles,

$$
\langle g_1 g_2 \cdots q_{n-1} \bar{q}_n | \mathcal{L}_{\text{hard}} | 0 \rangle^{\text{tree}}
$$
$$
= C_{+\cdots(--)}^{a_1 a_2 \cdots \alpha_{n-1} \bar{\alpha}_n}(\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_{n-1}, \tilde{p}_n) \, .
\tag{5.93}
$$

Here and below, $g_i \equiv g_\pm^{a_i}(p_i)$ stands for a gluon with helicity $\pm$, momentum p_i, color a_i, and analogously for (anti)quarks. From Eq. (5.93) we obtain the generic tree-level matching equation

$$
C_{+\cdots(--)}^{a_1 \cdots \bar{\alpha}_n}(\tilde{p}_1, \ldots, \tilde{p}_n) = -i \mathcal{T}^{\text{tree}}(g_1 \cdots \bar{q}_n) \, ,
\tag{5.94}
$$

where $\mathcal{I}^{\text{tree}}$ denotes the tree-level QCD helicity amplitude. Intuitively, since all external particles are energetic and well separated, we are away from any soft or collinear limits and so all propagators in the QCD tree-level diagram are far off shell and can be shrunk to a point. Hence, the tree-level diagram simply becomes the Wilson coefficient in SCET.

The above discussion can be extended to higher orders in perturbation theory. In pure dimensional regularization (where ϵ is used to simultaneously regulate UV and IR divergences) all bare loop graphs in SCET are scaleless and vanish. Here the UV and IR divergences precisely cancel each other, and the bare matrix elements are given by their tree-level expressions, Eq. (5.93). Including the counterterm $\delta_O(\epsilon_{\text{UV}})$ due to operator renormalization removes the UV divergences and leaves the IR divergences. Schematically, the renormalized loop amplitude computed in SCET using $\mathcal{L}_{\text{hard}}$ is

$$\mathcal{I}_{\text{SCET}} = \int (\langle \vec{O}^\dagger \rangle^{\text{tree}} + \langle \vec{O}^\dagger \rangle^{\text{loop}}) \, i\vec{C} = \big[1 + \delta_O(\epsilon_{\text{IR}}) \big] i\vec{C}, \qquad (5.95)$$

where we used that the loop contribution is a pure counterterm and thus proportional to the tree-level expression. In general, the counterterm δ_O is a matrix in color space, as we will see explicitly in Sec. 5.8 and Appendix B.7. By construction, the $1/\epsilon$ IR divergences in the effective theory, $C\,\delta_O(\epsilon_{\text{IR}})$, have to exactly match those of the full theory. Therefore, beyond tree level the matching coefficients in $\overline{\text{MS}}$ are given by the infrared-finite part of the renormalized full-theory amplitude, $\mathcal{A}_{\text{ren}}$, computed in pure dimensional regularization. The IR-finite part is obtained by multiplying $\mathcal{A}_{\text{ren}}$ by SCET $\overline{\text{MS}}$ renormalization factors, which cancel the full theory $1/\epsilon_{\text{IR}}$ poles. Decomposing the renormalized QCD amplitude in a color basis so that $\mathcal{A}_{\text{ren}}^{a_1 \cdots \alpha_n} = \bar{T}^{a_1 \cdots \bar{\alpha}_n} \vec{\mathcal{A}}_{\text{ren}}(g_1 \cdots \bar{q}_n)$, the all-orders form of Eq. (5.94) becomes

$$C_{+\cdots(\cdots-)}^{'a_1 \cdots \bar{\alpha}_n}(\tilde{p}_1, \ldots, \tilde{p}_n) = -i\mathcal{I}_{\text{fin}}(g_1 \cdots \bar{q}_n) \qquad (5.96)$$

$$\equiv \frac{-i\,\bar{T}^{a_1 \cdots \bar{\alpha}_n} \widehat{Z}_C^{-1} \vec{\mathcal{A}}_{\text{ren}}(g_1 \cdots \bar{q}_n)}{Z_\xi^{n_q/2} Z_A^{n_g/2}}.$$

The SCET renormalization factors $\widehat{Z}_C$, Z_ξ, and Z_A are discussed in Sec. 5.8.1. At one-

251

loop order this corresponds to taking $(-i\mathcal{A}_{\text{ren}}^{a_1 \cdots \bar{\alpha}_n})$ and simply dropping the $1/\epsilon_{\text{IR}}$ terms. In Sec. 5.4.2 we will discuss in more detail the use of different renormalization schemes to compute $\vec{\mathcal{A}}_{\text{ren}}(g_1 \cdots \bar{q}_n)$.

If the same color decomposition is used for the QCD amplitude as for the Wilson coefficients in Eq. (5.52), we can immediately read off the coefficients $\vec{C}$ in this color basis from Eq. (5.96). As an example, consider for simplicity the leading color n gluon amplitude, which has the color decomposition (see Appendix B.1.3)

$$\mathcal{I}_n(g_1 \cdots g_n) = ig_s^{n-2} \sum_{\sigma \in S_n/Z_n} \text{tr}[T^{a_{\sigma(1)}} \cdots T^{a_{\sigma(n)}}]$$
$$\times \sum_i g_s^i A_n^{(i)}(\sigma(1), \ldots, \sigma(n)), \tag{5.97}$$

where the first sum runs over all permutations σ of n objects (S_n) excluding cyclic permutations (Z_n). The $A_n^{(i)}$ are the color-ordered or partial amplitudes at i loops. Each is separately gauge invariant and only depends on the external momenta and helicities $(p_i\pm) \equiv (i^\pm)$. If we choose

$$T_k^{a_1 \cdots a_n} = \text{tr}[T^{a_{\sigma_k(1)}} \cdots T^{a_{\sigma_k(n)}}], \tag{5.98}$$

as the color basis in Eq. (5.52), where σ_k is the kth permutation in S_n/Z_n, then the Wilson coefficients in this color basis are given directly by

$$C_{\lambda_1 \cdots \lambda_2}^k(\tilde{p}_1, \ldots, \tilde{p}_n) =$$
$$g_s^{n-2} \sum_i g_s^i A_{n,\text{fin}}^{(i)}(\sigma_k(1^{\lambda_1}), \ldots, \sigma_k(n^{\lambda_2})), \tag{5.99}$$

where the subscript "fin" denotes the IR-finite part of the helicity amplitude, as defined in Eq. (5.96). This is easily extended beyond leading color, given a valid choice of subleading color basis. Our basis therefore achieves seamless matching from QCD helicity amplitudes onto SCET operators.

252

5.4.2 Renormalization Schemes

In this section we discuss in more detail the issue of renormalization/regularization schemes in QCD and in SCET. In particular, the construction of a basis of helicity operators discussed in Sec. 5.3 relied heavily on massless quarks and gluons having two helicity states, which is a feature specific to 4 dimensions. We clarify this issue here and discuss the conversion between various schemes.

In dimensional regularization, divergences are regularized by analytically continuing the particle momenta to d dimensions. In a general scheme, the helicities of quarks and gluons live in d_s^g, d_s^q dimensional spaces respectively. We shall here restrict ourselves to schemes where quarks have two helicities, but d_s^g is analytically continued. This is true of most commonly used regularization schemes, but is not necessary [337]. Different schemes within dimensional regularization differ in their treatment of d_s^g for internal (unobserved) and external (observed) particles. In the conventional dimensional regularization (CDR), 't Hooft-Veltman (HV) [338], and four-dimensional helicity (FDH) [339, 340] schemes the internal/external polarizations are treated in d/d (CDR), $d/4$ (HV), $4/4$ (FDH) dimensions.

For helicity-based computations, the FDH scheme has the advantage of having all helicities defined in 4 dimensions, where the spinor-helicity formalism applies, as well as preserving supersymmetry. Indeed, most of the recent one-loop computations of helicity amplitudes utilize on-shell methods and therefore employ the FDH scheme. However, most existing calculations of SCET matrix elements (jet, beam, and soft functions) use d-dimensional internal gluons, corresponding to the CDR/HV schemes.[8] As we will discuss below, CDR and HV are identical for matching onto SCET.

Although the FDH scheme is convenient for helicity amplitude computations, it leads to subtleties beyond NLO [343, 344]. As explained in Ref. [344], this discrepancy arises due to the different number of dimensions for the momenta in the loop integral and the spin space, leading to components of the gluon field whose couplings to quarks are not protected

---

[8]Recently while this chapter was being finalized, a calculation of the inclusive jet and soft functions in both FDH and dimensional reduction (DRED) [341] appeared in Ref. [342]. The conclusions of this section agree with their study of the regularization scheme dependence of QCD amplitudes.

by gauge invariance and require separate renormalization. Nevertheless, it has been shown that FDH is a consistent regularization scheme to NNLO [342]. The presence of these extra degrees of freedom in the FDH scheme is quite inconvenient in the formal construction of SCET, especially when working to subleading power. Because of this fact, and because most SCET calculations are performed in CDR/HV, our discussion of SCET schemes will focus on regularization schemes where the dimension of the gluon field and the momentum space are analytically continued in the same manner. We will also discuss how full-theory helicity amplitudes in the FDH scheme are converted to CDR/HV for the purposes of matching to SCET.

We will now describe how helicity amplitudes in the FDH scheme can be converted to CDR/HV. To get a finite correction from the $\mathcal{O}(\epsilon)$ part of the gluon polarization requires a factor from either ultraviolet (UV) or infrared (IR) $1/\epsilon$ divergences. Although the regularization of UV and IR divergences is coupled in pure dimensional regularization schemes by use of a common ϵ, they can in principle be separately regulated, and we discuss their role in the scheme conversion separately below.

When matching to SCET, the UV regulators in the full and effective theory need not be equal. Indeed, the effective theory does not reproduce the UV of the full theory. In massless QCD, scheme dependence due to the UV divergences only affects the coupling constant through virtual (internal) gluons. Therefore, the CDR and HV schemes have the same standard $\overline{\text{MS}}$ coupling, $\alpha_s(\mu)$, while FDH has a different coupling, $\alpha_s^{\text{FDH}}(\mu)$. The conversion between these couplings is achieved by a perturbatively calculable shift, known to two loops [345, 244, 340]

$$
\alpha_s^{\text{FDH}}(\mu) = \alpha_s(\mu) \left[1 + \frac{C_A}{3} \frac{\alpha_s(\mu)}{4\pi} \right.
$$
$$
\left. + \left(\frac{22}{9} C_A^2 - 2 C_F T_F n_f \right) \left(\frac{\alpha_s(\mu)}{4\pi} \right)^2 \right].
\tag{5.100}
$$

This replacement rule for the coupling captures the effect of the scheme choice from UV divergences. One can therefore perform a matching calculation, treating α_s in the full and

effective theories as independent parameters that can be defined in different schemes. A conversion between schemes can then be used to ensure that the matching coefficients are written entirely in terms of α_s defined in one scheme, for example using Eq. (5.100). The issue of UV regularization is therefore simple to handle in the matching.

The structure of $1/\epsilon^2$ and $1/\epsilon$ IR divergences in one-loop QCD amplitudes is well known, and allows one to determine their effect on converting amplitudes from FDH to CDR/HV. For a QCD amplitude involving n_q (anti)quarks and n_g gluons the FDH and HV one-loop amplitudes $\mathcal{I}^{(1)}$ are related by [244, 337]

$$\mathcal{I}^{(1)}_{\text{HV}} = \mathcal{I}^{(1)}_{\text{FDH}} - \frac{\alpha_s}{4\pi}\Big(\frac{n_q}{2}C_F + \frac{n_g}{6}C_A\Big)\mathcal{I}^{(0)}\,, \tag{5.101}$$

where $\mathcal{I}^{(0)}$ denotes the tree-level amplitude, and the precise scheme of the α_s entering here is a two-loop effect. At one loop, the FDH scheme can therefore be consistently used when calculating full-theory helicity amplitudes and results can easily be converted to HV with Eqs. (5.100) and (5.101) for use in SCET Wilson coefficients.

We will now compare CDR and HV schemes for SCET calculations and the construction of the operator basis. In the HV scheme, all external polarizations are 4 dimensional, so that one can use a basis of helicity operators, as was constructed in Sec. 5.3. However, in CDR external polarizations are d dimensional, with the limit $d \to 4$ taken. In particular, this implies that one must work with $d - 2$ gluon polarizations at intermediate steps, potentially allowing for the presence of evanescent operators corresponding to operators involving the additional components of the gluon field, so-called ϵ-helicities. However, we will now argue that there is no real distinction between the two schemes, and that one does not need to consider evanescent operators in SCET at leading power.

First consider the Wilson coefficients and matching. In the case of CDR, the operator basis must be extended to include operators involving the ϵ-helicities. However, their presence does not affect the matching coefficients for operators with physical helicities, since they do not contribute at tree level and all loop corrections are scaleless and vanish. Additionally, in Sec. 5.8, we will discuss the fact that the SCET renormalization of the operators is spin

independent at leading power, and therefore there is no mixing under renormalization group evolution between the physical and evanescent operators. For the beam and jet functions, azimuthal symmetry implies that the difference between a field with 2 or $2 - 2\epsilon$ polarizations is simply an overall factor of $1 - \epsilon$ and thus can be easily taken into account. The independence of the soft function to the differences in the CDR/HV regularization schemes follows from the insensitivity of the soft emission to the polarization of the radiating parton, which is made manifest by the SCET Lagrangian and the fact that the soft function can be written as a matrix element of Wilson lines. Thus there is no difference between CDR and HV and the helicity operator basis suffices.

5.5 Higgs + Jets

In this section, we consider the production of an on-shell Higgs + jets. We give the helicity operator basis and matching relations for $H + 0, 1, 2$ jets, and the corresponding helicity amplitudes are collected in Appendix B.2.

5.5.1 $H + 0$ Jets

The ggH and $q\bar{q}H$ processes contribute to the $H + 0$ jets process. For $q\bar{q}H$, the scalar current in Eq. (5.38) is required, and the helicity operator basis is given by

$$
\begin{aligned}
O_1^{\bar{\alpha}\beta} &= J_{12\,0}^{\bar{\alpha}\beta}\, H_3\,, \\
O_2^{\bar{\alpha}\beta} &= (J^\dagger)_{12\,0}^{\bar{\alpha}\beta}\, H_3\,,
\end{aligned}
\tag{5.102}
$$

with the unique color structure

$$
\bar{T}^{\alpha\bar{\beta}} = \left(\delta_{\alpha\bar{\beta}} \right)\,.
\tag{5.103}
$$

These operators are relevant when considering Higgs decays to massive quarks, for example $H \to \bar{b}b$. However, we will not consider this case further since for Higgs production the $b\bar{b}H$ and $t\bar{t}H$ contributions are much smaller than the dominant gluon-fusion hard scattering

256

process.

For ggH, the basis of helicity operators is given by

$$O_{++}^{ab} = \frac{1}{2}\,\mathcal{B}_{1+}^{a}\,\mathcal{B}_{2+}^{b}\,H_{3}\,,$$
$$O_{--}^{ab} = \frac{1}{2}\,\mathcal{B}_{1-}^{a}\,\mathcal{B}_{2-}^{b}\,H_{3}\,. \tag{5.104}$$

The operator O_{+-} is not allowed by angular momentum conservation. Similar helicity operators, extended to include the decay of the Higgs, were used in Ref. [14]. There is again a unique color structure for this process,

$$\bar{T}^{ab} = \left(\delta^{ab}\right)\,. \tag{5.105}$$

Writing the QCD helicity amplitudes as

$$\mathcal{I}(g_{1}g_{2}H_{3}) = \mathrm{i}\delta^{a_{1}a_{2}}\,A(1,2;3_{H})\,, \tag{5.106}$$

the Wilson coefficients for ggH are given by

$$\vec{C}_{++}(\tilde{p}_{1},\tilde{p}_{2};\tilde{p}_{3}) = A_{\mathrm{fin}}(1^{+},2^{+};3_{H})\,,$$
$$\vec{C}_{--}(\tilde{p}_{1},\tilde{p}_{2};\tilde{p}_{3}) = A_{\mathrm{fin}}(1^{-},2^{-};3_{H})\,. \tag{5.107}$$

The subscript "fin" in Eq. (5.107) denotes the IR-finite part of the helicity amplitudes, as discussed in Sec. 5.4. Note that the two amplitudes appearing in Eq. (5.107) are related by parity. The results for the gluon amplitudes up to NNLO are given in Appendix B.2.1. They correspond to the usual gluon-fusion process, where the Higgs couples to a (top) quark loop at leading order. The LO amplitude including the dependence on the mass of the quark running in the loop is well known. The NLO amplitudes are also known including the full quark-mass dependence [346, 347, 348, 349, 350], while the NNLO [351, 352, 353] and N^{3}LO [280, 281] amplitudes are known in an expansion in m_{H}/m_{t}.

5.5.2 $H + 1$ Jet

The $gq\bar{q}H$ and $gggH$ processes contribute to the $H + 1$ jet process. For $gq\bar{q}$, the basis of helicity operators is given by

$$
\begin{aligned}
O^{a\,\bar{\alpha}\beta}_{+(+)} &= \mathcal{B}^a_{1+}\, J^{\bar{\alpha}\beta}_{23+}\, H_4\,, \\
O^{a\,\bar{\alpha}\beta}_{-(+)} &= \mathcal{B}^a_{1-}\, J^{\bar{\alpha}\beta}_{23+}\, H_4\,, \\
O^{a\,\bar{\alpha}\beta}_{+(-)} &= \mathcal{B}^a_{1+}\, J^{\bar{\alpha}\beta}_{23-}\, H_4\,, \\
O^{a\,\bar{\alpha}\beta}_{-(-)} &= \mathcal{B}^a_{1-}\, J^{\bar{\alpha}\beta}_{23-}\, H_4\,.
\end{aligned}
\tag{5.108}
$$

Note that we consider only QCD corrections to the ggH process, so the $q\bar{q}$ pair is described by $J_{ij\pm}$. For ggg, the helicity operator basis is

$$
\begin{aligned}
O^{abc}_{+++} &= \frac{1}{3!}\, \mathcal{B}^a_{1+}\, \mathcal{B}^b_{2+}\, \mathcal{B}^c_{3+}\, H_4\,, \\
O^{abc}_{++-} &= \frac{1}{2}\, \mathcal{B}^a_{1+}\, \mathcal{B}^b_{2+}\, \mathcal{B}^c_{3-}\, H_4\,, \\
O^{abc}_{--+} &= \frac{1}{2}\, \mathcal{B}^a_{1-}\, \mathcal{B}^b_{2-}\, \mathcal{B}^c_{3+}\, H_4\,, \\
O^{abc}_{---} &= \frac{1}{3!}\, \mathcal{B}^a_{1-}\, \mathcal{B}^b_{2-}\, \mathcal{B}^c_{3-}\, H_4\,.
\end{aligned}
\tag{5.109}
$$

For both cases the color space is one dimensional and we use the respective color structures as basis elements

$$
\bar{T}^{a\alpha\bar{\beta}} = \left(T^a_{\alpha\bar{\beta}} \right)\,, \qquad \bar{T}^{abc} = \left(\mathrm{i} f^{abc} \right)\,.
\tag{5.110}
$$

In principle, there could be another independent color structure, d^{abc}, for $gggH$. The $gggH$ operators transform under charge conjugation as

$$
\begin{aligned}
\mathrm{C}\, O^{abc}_{\lambda_1\lambda_2\lambda_3}&(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3;\tilde{p}_4)\, \bar{T}^{abc}\, \mathrm{C} \\
&= -O^{cba}_{\lambda_1\lambda_2\lambda_3}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3;\tilde{p}_4)\, \bar{T}^{abc}\,.
\end{aligned}
\tag{5.111}
$$

258

Charge conjugation invariance of QCD thus leads to

$$C^{abc}_{\lambda_1\lambda_2\lambda_3}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3;\tilde{p}_4) = -C^{cba}_{\lambda_1\lambda_2\lambda_3}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3;\tilde{p}_4)\,, \tag{5.112}$$

which implies that the d^{abc} color structure cannot arise to all orders in perturbation theory, so it suffices to consider $\mathrm{i}f^{abc}$ as in Eq. (5.110). This also means that the d^{abc} color structure cannot be generated by mixing under renormalization group evolution, which will be seen explicitly in Eq. (5.259).

Using Eq. (5.110), we write the QCD helicity amplitudes as

$$\mathcal{I}(g_1 g_2 g_3 H_4) = \mathrm{i}\,(\mathrm{i}f^{a_1 a_2 a_3})\,A(1,2,3;4_H)\,,$$
$$\mathcal{I}(g_1 q_2 \bar{q}_3 H_4) = \mathrm{i}\,T^{a_1}_{\alpha_2\bar{\alpha}_3}\,A(1;2_q,3_{\bar{q}};4_H)\,. \tag{5.113}$$

The Wilson coefficients for $gq\bar{q}H$ are then given by

$$\vec{C}_{+(+)}(\tilde{p}_1;\tilde{p}_2,\tilde{p}_3;\tilde{p}_4) = A_{\mathrm{fin}}(1^+;2^+_q,3^-_{\bar{q}};4_H)\,,$$
$$\vec{C}_{-(+)}(\tilde{p}_1;\tilde{p}_2,\tilde{p}_3;\tilde{p}_4) = A_{\mathrm{fin}}(1^-;2^+_q,3^-_{\bar{q}};4_H)\,,$$
$$\vec{C}_{+(-)}(\tilde{p}_1;\tilde{p}_2,\tilde{p}_3;\tilde{p}_4) = \vec{C}_{+(+)}(\tilde{p}_1;\tilde{p}_3,\tilde{p}_2;\tilde{p}_4)\,,$$
$$\vec{C}_{-(-)}(\tilde{p}_1;\tilde{p}_2,\tilde{p}_3;\tilde{p}_4) = \vec{C}_{-(+)}(\tilde{p}_1;\tilde{p}_3,\tilde{p}_2;\tilde{p}_4)\,, \tag{5.114}$$

where the last two coefficients follow from charge conjugation invariance. The Wilson coefficients for $gggH$ are given by

$$\vec{C}_{+++}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3;\tilde{p}_4) = A_{\mathrm{fin}}(1^+,2^+,3^+;4_H)\,,$$
$$\vec{C}_{++-}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3;\tilde{p}_4) = A_{\mathrm{fin}}(1^+,2^+,3^-;4_H)\,,$$
$$\vec{C}_{--+}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3;\tilde{p}_4) = \vec{C}_{++-}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3;\tilde{p}_4)\Big|_{\langle..\rangle\leftrightarrow[..]}\,,$$
$$\vec{C}_{---}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3;\tilde{p}_4) = \vec{C}_{+++}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3;\tilde{p}_4)\Big|_{\langle..\rangle\leftrightarrow[..]}\,, \tag{5.115}$$

259

where the last two relations follow from parity invariance. As before, the subscript "fin" in Eqs. (5.114) and (5.115) denotes the finite part of the IR divergent amplitudes. The NLO helicity amplitudes were calculated in Ref. [252], and are given in Appendix B.2.2, and the NNLO helicity amplitudes were calculated in Ref. [354]. Both calculations were performed in the $m_t \to \infty$ limit. At NLO, the first corrections in m_H^2/m_t^2 were obtained in Ref. [355].

5.5.3 $H + 2$ Jets

For $H + 2$ jets, the $q\bar{q}\,q'\bar{q}'H$, $q\bar{q}\,q\bar{q}H$, $ggq\bar{q}H$, and $ggggH$ processes contribute, each of which we discuss in turn. Again, we consider only QCD corrections to the ggH process, so $q\bar{q}$ pairs are described by the helicity currents $J_{ij\pm}$. The LO helicity amplitudes for $H + 2$ jets in the $m_t \to \infty$ limit were calculated in Refs. [356, 251] and are collected in Appendix B.2.3 for each channel. The LO amplitudes including the m_t dependence were calculated in [357] (but explicit expressions for $ggggH$ were not given due to their length). The NLO helicity amplitudes were computed in Refs. [358, 359, 360, 253, 254, 361].

$q\bar{q}\,q'\bar{q}'H$ and $q\bar{q}\,q\bar{q}H$

For the case of distinct quark flavors, $q\bar{q}\,q'\bar{q}'H$, the helicity basis consists of four independent operators,

$$
\begin{aligned}
O_{(+;+)}^{\bar{\alpha}\beta\bar{\gamma}\delta} &= J_{q\,12+}^{\bar{\alpha}\beta}\, J_{q'\,34+}^{\bar{\gamma}\delta}\, H_5\,, \\
O_{(+;-)}^{\bar{\alpha}\beta\bar{\gamma}\delta} &= J_{q\,12+}^{\bar{\alpha}\beta}\, J_{q'\,34-}^{\bar{\gamma}\delta}\, H_5\,, \\
O_{(-;+)}^{\bar{\alpha}\beta\bar{\gamma}\delta} &= J_{q\,12-}^{\bar{\alpha}\beta}\, J_{q'\,34+}^{\bar{\gamma}\delta}\, H_5\,, \\
O_{(-;-)}^{\bar{\alpha}\beta\bar{\gamma}\delta} &= J_{q\,12-}^{\bar{\alpha}\beta}\, J_{q'\,34-}^{\bar{\gamma}\delta}\, H_5\,,
\end{aligned}
\tag{5.116}
$$

where the additional labels on the quark currents indicate the quark flavors. For the case of identical quark flavors, $q\bar{q}\,q\bar{q}H$, the basis only has three independent helicity operators,

$$
O_{(++)}^{\bar{\alpha}\beta\bar{\gamma}\delta} = \frac{1}{4}\, J_{12+}^{\bar{\alpha}\beta}\, J_{34+}^{\bar{\gamma}\delta}\, H_5\,,
$$

$$O_{(+-)}^{\bar{\alpha}\beta\bar{\gamma}\delta} = J_{12+}^{\bar{\alpha}\beta} \, J_{34-}^{\bar{\gamma}\delta} \, H_5 \,,$$

$$O_{(--)}^{\bar{\alpha}\beta\bar{\gamma}\delta} = \frac{1}{4} \, J_{12-}^{\bar{\alpha}\beta} \, J_{34-}^{\bar{\gamma}\delta} \, H_5 \,,\tag{5.117}$$

since both quark currents have the same flavor. In both cases we use the color basis

$$\bar{T}^{\alpha\bar{\beta}\gamma\bar{\delta}} = 2T_F \left(\delta_{\alpha\bar{\delta}} \, \delta_{\gamma\bar{\beta}} \,, \; \delta_{\alpha\bar{\beta}} \, \delta_{\gamma\bar{\delta}} \right).\tag{5.118}$$

The QCD helicity amplitudes for $q\bar{q}\,q'\bar{q}'H$ can be color decomposed in the basis of Eq. (5.118) as

$$\mathcal{I}(q_1\bar{q}_2 q_3'\bar{q}_4' H_5) = 2\mathrm{i}T_F \Big[\delta_{\alpha_1\bar{\alpha}_4}\delta_{\alpha_3\bar{\alpha}_2} A(1_q, 2_{\bar{q}}; 3_{q'}, 4_{\bar{q}'}; 5_H)$$

$$+ \frac{1}{N} \, \delta_{\alpha_1\bar{\alpha}_2}\delta_{\alpha_3\bar{\alpha}_4} B(1_q, 2_{\bar{q}}; 3_{q'}, 4_{\bar{q}'}; 5_H) \Big] \,,\tag{5.119}$$

where we have included a factor of $1/N$ for convenience. The amplitude vanishes when the quark and antiquark of the same flavor have the same helicity, in accordance with the fact that the operators of Eq. (5.116) provide a complete basis of helicity operators. For identical quark flavors, $q\bar{q}\,q\bar{q}H$, the amplitudes can be obtained from the $q\bar{q}\,q'\bar{q}'H$ amplitudes using the relation

$$\mathcal{I}(q_1\bar{q}_2 q_3\bar{q}_4 H_5) = \mathcal{I}(q_1\bar{q}_2 q_3'\bar{q}_4' H_5) - \mathcal{I}(q_1\bar{q}_4 q_3'\bar{q}_2' H_5) \,.\tag{5.120}$$

The Wilson coefficients for $q\bar{q}\,q'\bar{q}'H$ are then given by

$$\vec{C}_{(+;+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5) = \begin{pmatrix} A_{\mathrm{fin}}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-; 5_H) \\ \frac{1}{N} B_{\mathrm{fin}}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-; 5_H) \end{pmatrix} \,,$$

$$\vec{C}_{(+;-)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5) = \begin{pmatrix} A_{\mathrm{fin}}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+; 5_H) \\ \frac{1}{N} B_{\mathrm{fin}}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+; 5_H) \end{pmatrix} \,,$$

$$\vec{C}_{(-;+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5) = \vec{C}_{(+;-)}(\tilde{p}_2, \tilde{p}_1; \tilde{p}_4, \tilde{p}_3; \tilde{p}_5) \,,$$

$$\vec{C}_{(-;-)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5) = \vec{C}_{(+;+)}(\tilde{p}_2, \tilde{p}_1; \tilde{p}_4, \tilde{p}_3; \tilde{p}_5) \,,\tag{5.121}$$

261

and for $q\bar{q}\,q\bar{q}H$ they are given in terms of the amplitudes A_{fin} and B_{fin} for $q\bar{q}\,q'\bar{q}'H$ by

$$\vec{C}_{(++)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_5) = \begin{pmatrix} A_{\text{fin}}(1_q^+,2_{\bar{q}}^-;3_q^+,4_{\bar{q}}^-;5_H) - \frac{1}{N}B_{\text{fin}}(1_q^+,4_{\bar{q}}^-;3_q^+,2_{\bar{q}}^-;5_H) \\[2mm] \frac{1}{N}B_{\text{fin}}(1_q^+,2_{\bar{q}}^-;3_q^+,4_{\bar{q}}^-;5_H) - A_{\text{fin}}(1_q^+,4_{\bar{q}}^-;3_q^+,2_{\bar{q}}^-;5_H) \end{pmatrix},$$

$$\vec{C}_{(+-)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_5) = \begin{pmatrix} A_{\text{fin}}(1_q^+,2_{\bar{q}}^-;3_q^-,4_{\bar{q}}^+;5_H) \\[2mm] \frac{1}{N}B_{\text{fin}}(1_q^+,2_{\bar{q}}^-;3_q^-,4_{\bar{q}}^+;5_H) \end{pmatrix},$$

$$\vec{C}_{(--)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_5) = \vec{C}_{(++)}(\tilde{p}_2,\tilde{p}_1;\tilde{p}_4,\tilde{p}_3;\tilde{p}_5)\,. \tag{5.122}$$

The relations for $\vec{C}_{(-;\pm)}$ and $\vec{C}_{(--)}$ follow from charge conjugation invariance. Note that there is no exchange term for $\vec{C}_{(+-)}$, since the amplitude vanishes when the quark and antiquark of the same flavor have the same helicity (both $+$ or both $-$). Also, recall that the symmetry factors of $1/4$ in Eq. (5.117) already take care of the interchange of identical (anti)quarks, so there are no additional symmetry factors needed for $\vec{C}_{(++)}$. Explicit expressions for the required amplitudes at tree level are given in Appendix B.2.3.

$ggq\bar{q}H$

For $ggq\bar{q}H$, the helicity basis consists of a total of six independent operators,

$$O_{++(+)}^{ab\,\bar{\alpha}\beta} = \frac{1}{2}\,\mathcal{B}_{1+}^a\,\mathcal{B}_{2+}^b\,J_{34+}^{\bar{\alpha}\beta}\,H_5\,,$$

$$O_{+-(+)}^{ab\,\bar{\alpha}\beta} = \mathcal{B}_{1+}^a\,\mathcal{B}_{2-}^b\,J_{34+}^{\bar{\alpha}\beta}\,H_5\,,$$

$$O_{--(+)}^{ab\,\bar{\alpha}\beta} = \frac{1}{2}\mathcal{B}_{1-}^a\,\mathcal{B}_{2-}^b\,J_{34+}^{\bar{\alpha}\beta}\,H_5\,,$$

$$O_{++(-)}^{ab\,\bar{\alpha}\beta} = \frac{1}{2}\,\mathcal{B}_{1+}^a\,\mathcal{B}_{2+}^b\,J_{34-}^{\bar{\alpha}\beta}\,H_5\,,$$

$$O_{+-(-)}^{ab\,\bar{\alpha}\beta} = \mathcal{B}_{1+}^a\,\mathcal{B}_{2-}^b\,J_{34-}^{\bar{\alpha}\beta}\,H_5\,,$$

$$O_{--(-)}^{ab\,\bar{\alpha}\beta} = \frac{1}{2}\mathcal{B}_{1-}^a\,\mathcal{B}_{2-}^b\,J_{34-}^{\bar{\alpha}\beta}\,H_5\,. \tag{5.123}$$

We use the color basis already given in Eq. (5.61),

$$\bar{T}^{ab\alpha\bar{\beta}} = \left((T^a T^b)_{\alpha\bar{\beta}} \, , \, (T^b T^a)_{\alpha\bar{\beta}} \, , \, \mathrm{tr}[T^a T^b] \, \delta_{\alpha\bar{\beta}} \right) . \tag{5.124}$$

Using Eq. (5.124), the color decomposition of the QCD helicity amplitudes into partial amplitudes is

$$
\begin{aligned}
& \mathcal{I}\big(g_1 g_2 \, q_3 \bar{q}_4 H_5\big) \\
& = \mathrm{i} \sum_{\sigma \in S_2} \left[T^{a_{\sigma(1)}} T^{a_{\sigma(2)}} \right]_{\alpha_3 \bar{\alpha}_4} A(\sigma(1), \sigma(2); 3_q, 4_{\bar{q}}; 5_H) \\
& \quad + \mathrm{i} \, \mathrm{tr}[T^{a_1} T^{a_2}] \, \delta_{\alpha_3 \bar{\alpha}_4} \, B(1, 2; 3_q, 4_{\bar{q}}; 5_H) .
\end{aligned}
\tag{5.125}
$$

The B amplitudes vanish at tree level. From Eq. (5.125) we can read off the Wilson coefficients,

$$
\vec{C}_{+-(+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5) =
\begin{pmatrix}
A_{\mathrm{fin}}(1^+, 2^-; 3_q^+, 4_{\bar{q}}^-; 5_H) \\
A_{\mathrm{fin}}(2^-, 1^+; 3_q^+, 4_{\bar{q}}^-; 5_H) \\
B_{\mathrm{fin}}(1^+, 2^-; 3_q^+, 4_{\bar{q}}^-; 5_H)
\end{pmatrix} ,
$$

$$
\vec{C}_{++(+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5) =
\begin{pmatrix}
A_{\mathrm{fin}}(1^+, 2^+; 3_q^+, 4_{\bar{q}}^-; 5_H) \\
A_{\mathrm{fin}}(2^+, 1^+; 3_q^+, 4_{\bar{q}}^-; 5_H) \\
B_{\mathrm{fin}}(1^+, 2^+; 3_q^+, 4_{\bar{q}}^-; 5_H)
\end{pmatrix} ,
$$

$$
\vec{C}_{--(+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5) =
\begin{pmatrix}
A_{\mathrm{fin}}(1^-, 2^-; 3_q^+, 4_{\bar{q}}^-; 5_H) \\
A_{\mathrm{fin}}(2^-, 1^-; 3_q^+, 4_{\bar{q}}^-; 5_H) \\
B_{\mathrm{fin}}(1^-, 2^-; 3_q^+, 4_{\bar{q}}^-; 5_H)
\end{pmatrix} .
\tag{5.126}
$$

The Wilson coefficients of the last three operators in Eq. (5.123) are obtained by charge conjugation as discussed in Sec. 5.3.5. Under charge conjugation, the operators transform as

$$
\begin{aligned}
& C \, O^{ab\,\bar{\alpha}\beta}_{\lambda_1 \lambda_2(\pm)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5) \, \bar{T}^{ab\,\alpha\bar{\beta}} \, C \\
& = -O^{ba\,\bar{\alpha}\beta}_{\lambda_1 \lambda_2(\mp)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_4, \tilde{p}_3; \tilde{p}_5) \, \bar{T}^{ab\,\alpha\bar{\beta}} ,
\end{aligned}
\tag{5.127}
$$

so charge conjugation invariance of QCD implies

$$\vec{C}_{\lambda_1\lambda_2(-)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5) = \widehat{V}\vec{C}_{\lambda_1\lambda_2(+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_4, \tilde{p}_3; \tilde{p}_5)$$

$$\text{with} \quad \widehat{V} = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{5.128}$$

Explicit expressions for the required amplitudes at tree level are given in Appendix B.2.3.

$ggggH$

For $ggggH$, the helicity basis consists of five independent operators,

$$O^{abcd}_{++++} = \frac{1}{4!}\,\mathcal{B}^a_{1+}\mathcal{B}^b_{2+}\mathcal{B}^c_{3+}\mathcal{B}^d_{4+}\,H_5\,,$$

$$O^{abcd}_{+++-} = \frac{1}{3!}\,\mathcal{B}^a_{1+}\mathcal{B}^b_{2+}\mathcal{B}^c_{3+}\mathcal{B}^d_{4-}\,H_5\,,$$

$$O^{abcd}_{++--} = \frac{1}{4}\,\mathcal{B}^a_{1+}\mathcal{B}^b_{2+}\mathcal{B}^c_{3-}\mathcal{B}^d_{4-}\,H_5\,,$$

$$O^{abcd}_{---+} = \frac{1}{3!}\,\mathcal{B}^a_{1-}\mathcal{B}^b_{2-}\mathcal{B}^c_{3-}\mathcal{B}^d_{4+}\,H_5\,,$$

$$O^{abcd}_{----} = \frac{1}{4!}\,\mathcal{B}^a_{1-}\mathcal{B}^b_{2-}\mathcal{B}^c_{3-}\mathcal{B}^d_{4-}\,H_5\,. \tag{5.129}$$

We use the basis of color structures

$$\vec{T}^{abcd} = \frac{1}{2\cdot 2T_F}\begin{pmatrix} \text{tr}[abcd] + \text{tr}[dcba] \\ \text{tr}[acdb] + \text{tr}[bdca] \\ \text{tr}[adbc] + \text{tr}[cbda] \\ 2\text{tr}[ab]\,\text{tr}[cd] \\ 2\text{tr}[ac]\,\text{tr}[db] \\ 2\text{tr}[ad]\,\text{tr}[bc] \end{pmatrix}^T, \tag{5.130}$$

264

where we have used the shorthand notation

$$\mathrm{tr}[ab] = \mathrm{tr}[T^a T^b], \qquad \mathrm{tr}[abcd] = \mathrm{tr}[T^a T^b T^c T^d].$$ (5.131)

Note that the three independent color structures with a minus sign instead of the plus sign in the first three lines in Eq. (5.130) can be eliminated using charge conjugation invariance, see Sec. 5.7.1.

The color decomposition of the QCD helicity amplitudes into partial amplitudes using the color basis in Eq. (5.130) is

$$\mathcal{I}(g_1 g_2 g_3 g_4 H_5) = \frac{\mathrm{i}}{2T_F} \Bigg[\sum_{\sigma \in S_4/Z_4} \mathrm{tr}[a_{\sigma(1)} a_{\sigma(2)} a_{\sigma(3)} a_{\sigma(4)}]$$
$$\times A\big(\sigma(1), \sigma(2), \sigma(3), \sigma(4); 5_H\big)$$
$$+ \sum_{\sigma \in S_4/Z_2^3} \mathrm{tr}[a_{\sigma(1)} a_{\sigma(2)}] \mathrm{tr}[a_{\sigma(3)} a_{\sigma(4)}]$$
$$\times B\big(\sigma(1), \sigma(2), \sigma(3), \sigma(4); 5_H\big) \Bigg],$$ (5.132)

where the B amplitudes vanish at tree level. From Eq. (5.132) we obtain the Wilson coefficients,

$$\vec{C}_{++--}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3, \tilde{p}_4; \tilde{p}_5) = \begin{pmatrix} 2A_{\mathrm{fin}}(1^+, 2^+, 3^-, 4^-; 5_H) \\ 2A_{\mathrm{fin}}(1^+, 3^-, 4^-, 2^+; 5_H) \\ 2A_{\mathrm{fin}}(1^+, 4^-, 2^+, 3^-; 5_H) \\ B_{\mathrm{fin}}(1^+, 2^+, 3^-, 4^-; 5_H) \\ B_{\mathrm{fin}}(1^+, 3^-, 4^-, 2^+; 5_H) \\ B_{\mathrm{fin}}(1^+, 4^-, 2^+, 3^-; 5_H) \end{pmatrix},$$

$$\vec{C}_{+++-}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3, \tilde{p}_4; \tilde{p}_5) = \begin{pmatrix} 2A_{\text{fin}}(1^+, 2^+, 3^+, 4^-; 5_H) \\ 2A_{\text{fin}}(1^+, 3^+, 4^-, 2^+; 5_H) \\ 2A_{\text{fin}}(1^+, 4^-, 2^+, 3^+; 5_H) \\ B_{\text{fin}}(1^+, 2^+, 3^+, 4^-; 5_H) \\ B_{\text{fin}}(1^+, 3^+, 4^-, 2^+; 5_H) \\ B_{\text{fin}}(1^+, 4^-, 2^+, 3^+; 5_H) \end{pmatrix},$$

$$\vec{C}_{++++}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3, \tilde{p}_4; \tilde{p}_5) = \begin{pmatrix} 2A_{\text{fin}}(1^+, 2^+, 3^+, 4^+; 5_H) \\ 2A_{\text{fin}}(1^+, 3^+, 4^+, 2^+; 5_H) \\ 2A_{\text{fin}}(1^+, 4^+, 2^+, 3^+; 5_H) \\ B_{\text{fin}}(1^+, 2^+, 3^+, 4^+; 5_H) \\ B_{\text{fin}}(1^+, 3^+, 4^+, 2^+; 5_H) \\ B_{\text{fin}}(1^+, 4^+, 2^+, 3^+; 5_H) \end{pmatrix},$$

$$\vec{C}_{---+}(\tilde{p}_1, \ldots; \tilde{p}_5) = \vec{C}_{+++-}(\tilde{p}_1, \ldots; \tilde{p}_5)\Big|_{\langle ..\rangle \leftrightarrow [..]},$$

$$\vec{C}_{----}(\tilde{p}_1, \ldots; \tilde{p}_5) = \vec{C}_{++++}(\tilde{p}_1, \ldots; \tilde{p}_5)\Big|_{\langle ..\rangle \leftrightarrow [..]}. \tag{5.133}$$

The last two coefficients follow from parity invariance. The factors of two in the first three entries of the coefficients come from combining the two traces in the first three entries in Eq. (5.130) using charge conjugation invariance. Because of the cyclic symmetry of the traces, the partial amplitudes are invariant under the corresponding cyclic permutations of their first four arguments, which means that most of the amplitudes in Eq. (5.133) are not independent. Explicit expressions for the necessary amplitudes at tree level are given in Appendix B.2.3.

5.6 Vector Boson + Jets

In this section, we give the helicity operator basis and the corresponding matching for the production of a γ, Z, or W vector boson in association with up to two jets. The corresponding helicity amplitudes are collected in Appendix B.3.

We work at tree level in the electroweak coupling and consider only QCD corrections, so any external $q\bar{q}$ pairs are described by the helicity vector currents $J_{ij\pm}$ in Eq. (5.37). We always include the subsequent leptonic decays $\gamma/Z \to \ell\bar{\ell}$, $W^{\pm} \to \nu\bar{\ell}/\ell\bar{\nu}$. In the following, for γ/Z processes, ℓ stands for any charged lepton or neutrino flavor, and q stands for any quark flavor. For W processes, we use ℓ to denote any charged lepton flavor and ν the corresponding neutrino flavor. Similarly, we use u and d to denote any up-type or down-type quark flavor (i.e. not necessarily first generation quarks only).

The operators in the helicity bases satisfy the transformation properties under C and P as discussed in Sec. 5.3.5. However, the weak couplings in the amplitudes explicitly violate C and P. Therefore, to utilize the C and P transformations of the operators and minimize the number of required amplitudes and Wilson coefficients, it is useful to separate the weak couplings from the amplitudes.

We define P_Z and P_W as the ratios of the Z and W propagators to the photon propagator,

$$P_{Z,W}(s) = \frac{s}{s - m_{Z,W}^2 + i\Gamma_{Z,W}m_{Z,W}} \,. \tag{5.134}$$

The left- and right-handed couplings $v_{L,R}$ of a particle to the Z boson are, as usual,

$$v_L^i = \frac{2T_3^i - 2Q^i \sin^2\theta_W}{\sin(2\theta_W)} \,, \quad v_R^i = -\frac{2Q^i \sin^2\theta_W}{\sin(2\theta_W)} \,, \tag{5.135}$$

where T_3^i is the third component of weak isospin, Q^i is the electromagnetic charge in units of $|e|$, and θ_W is the weak mixing angle.

The γ/Z amplitudes can then be decomposed as

$$
\begin{aligned}
\mathcal{I}(\cdots \ell\bar{\ell}) \\
= e^2 \bigg\{ & \big[Q^\ell Q^q + v_{L,R}^\ell v_{L,R}^q P_Z(s_{\ell\bar{\ell}}) \big] \mathcal{I}_q(\cdots \ell\bar{\ell}) \\
& + \sum_{i=1}^{n_f} \Big[Q^\ell Q^i + v_{L,R}^\ell \frac{v_L^i + v_R^i}{2} P_Z(s_{\ell\bar{\ell}}) \Big] \mathcal{I}_v(\cdots \ell\bar{\ell})
\end{aligned}
$$

$$+ \frac{v_{L,R}^{\ell}}{\sin(2\theta_W)} P_Z(s_{\ell\bar{\ell}}) \mathcal{I}_a(\cdots \ell\bar{\ell}) \Bigg\} . \tag{5.136}$$

Here, $\mathcal{I}_q$ corresponds to the usual contribution where the vector boson couples directly to the external quark line with flavor q. (There is one such contribution for each external $q\bar{q}$ pair, and this contribution is absent for pure gluonic amplitudes like $gggZ$.) For $\mathcal{I}_v$, the γ/Z couples to an internal quark loop through a vector current and the sum runs over all considered internal quark flavors. For $\mathcal{I}_a$, the Z boson couples to an internal quark loop through the axial-vector current. This means that when using parity and charge conjugation we have to include an additional relative minus sign for this contribution. We have also made the assumption in Eq. (5.136) that all quarks, except for the top, are massless. Since $\mathcal{I}_a$ vanishes when summed over a massless isodoublet, this has the consequence that only the b, t isodoublet contributes to $\mathcal{I}_a$, hence the lack of sum over flavors. We have made this simplification following the one-loop calculation of Ref. [235], which calculated the amplitude in an expansion in $1/m_t^2$, assuming all other kinematic invariants to be smaller than the top mass. From the point of view of constructing a basis these assumptions are trivial to relax.

The $W^{\mp}$ amplitudes can be written as

$$\mathcal{I}(\cdots \ell^- \bar{\nu}^+) = \frac{e^2 V_{ud}}{2\sin^2\theta_W} P_W(s_{\ell\bar{\nu}}) \mathcal{I}_q(\cdots \ell^- \bar{\nu}^+) ,$$

$$\mathcal{I}(\cdots \nu^- \bar{\ell}^+) = \frac{e^2 V_{ud}^{\dagger}}{2\sin^2\theta_W} P_W(s_{\nu\bar{\ell}}) \mathcal{I}_q(\cdots \nu^- \bar{\ell}^+) , \tag{5.137}$$

where V_{ud} is the appropriate CKM-matrix element. The $\mathcal{I}_q$ amplitudes are the same in Eqs. (5.136) and (5.137), since all electroweak couplings have been extracted, but we have explicitly included the helicity labels (not to be mistaken as charge labels) to emphasize that these are the only possible helicities. The analogs of $\mathcal{I}_v$ and $\mathcal{I}_a$ do not exist for W production.

We note again that Eqs. (5.136) and (5.137) hold at tree level in the electroweak coupling, which is what we consider in this chapter. At this level, the leptons always couple to the

vector boson through the currents [see Eq. (B.19)]

$$\langle p_\ell \pm | \gamma^\mu | p_{\bar\ell} \pm \rangle = \langle p_{\bar\ell} \mp | \gamma^\mu | p_\ell \mp \rangle \,. \tag{5.138}$$

This allows us to obtain the Wilson coefficients for opposite lepton helicities simply by interchanging the lepton momenta.

5.6.1 $V + 0$ Jets

For $\gamma/Z + 0$ jets, the partonic process is $q\bar{q}\ell\bar{\ell}$, and the basis of helicity operators is

$$
\begin{aligned}
O_{(+;\pm)}^{\bar\alpha\beta} &= J_{q\,12+}^{\bar\alpha\beta}\, J_{\ell\,34\pm}\,, \\
O_{(-;\pm)}^{\bar\alpha\beta} &= J_{q\,12-}^{\bar\alpha\beta}\, J_{\ell\,34\pm}\,.
\end{aligned}
\tag{5.139}
$$

In principle, the process $gg\ell\bar{\ell}$ is allowed through the axial anomaly, but its contribution vanishes because in the matching calculation the gluons are taken to be on shell, and we neglect lepton masses.

For $W^\mp + 0$ jets, the partonic processes are $u\bar{d}\ell\bar{\nu}$ and $d\bar{u}\nu\bar{\ell}$, respectively. Since the W only couples to left-handed fields, the helicity basis simplifies to

$$
\begin{aligned}
O_{(W^-)}^{\bar\alpha\beta} &= J_{\bar{u}d\,12-}^{\bar\alpha\beta}\, J_{\bar{\ell}\nu\,34-}\,,. \\
O_{(W^+)}^{\bar\alpha\beta} &= J_{\bar{d}u\,12-}^{\bar\alpha\beta}\, J_{\bar{\nu}\ell\,34-}\,.
\end{aligned}
\tag{5.140}
$$

Here, we have explicitly written out the flavor structure of the currents. However, we use the shorthand subscript $(W^\mp)$ on the operators and Wilson coefficients, since we will not focus any further on the flavor structure. In an explicit calculation, one must of course sum over all relevant flavor combinations.

The unique color structure for $V + 0$ jets is

$$\bar{T}^{\alpha\bar\beta} = \left(\delta_{\alpha\bar\beta} \right)\,, \tag{5.141}$$

269

and extracting it from the amplitudes, we have

$$\mathcal{I}_{q,v,a}(q_1 \bar{q}_2 \ell_3 \bar{\ell}_4) = \mathrm{i}\,\delta_{\alpha_1 \bar{\alpha}_2}\,A_{q,v,a}(1_q, 2_{\bar{q}}; 3_\ell, 4_{\bar{\ell}})\,.\tag{5.142}$$

Here, A_v and A_a first appear at two loops. In addition, A_a is proportional to the top and bottom mass splitting due to isodoublet cancellations. It drops out when both top and bottom are treated as massless (e.g., when the matching scale is much larger than the top mass).

We use the same electroweak decomposition as in Eqs. (5.136) and (5.137) to write the Wilson coefficients. For $\gamma/Z + 0$ jets, we have

$$
\begin{aligned}
\vec{C}_{(\lambda_q; \lambda_\ell)}&(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4)\\
&= e^2 \Bigg\{ \big[Q^\ell Q^q + v_{\lambda_\ell}^\ell v_{\lambda_q}^q P_Z(s_{34}) \big] \vec{C}_{q(\lambda_q; \lambda_\ell)}(\ldots)\\
&\quad + \sum_{i=1}^{n_f} \Big[Q^\ell Q^i + v_{\lambda_\ell}^\ell \frac{v_L^i + v_R^i}{2} P_Z(s_{34}) \Big] \vec{C}_{v(\lambda_q; \lambda_\ell)}(\ldots)\\
&\quad + \frac{v_{\lambda_\ell}^\ell}{\sin(2\theta_W)} P_Z(s_{34})\, \vec{C}_{a(\lambda_q; \lambda_\ell)}(\ldots) \Bigg\}\,,
\end{aligned}\tag{5.143}
$$

where the weak couplings are determined by the helicity labels of the quark and lepton currents,

$$v_+^\ell = v_R^\ell\,, \quad v_-^\ell = v_L^\ell\,, \qquad v_+^q = v_R^q\,, \quad v_-^q = v_L^q\,.\tag{5.144}$$

For $W + 0$ jets, we simply have

$$
\begin{aligned}
\vec{C}_{(W-)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) &= \frac{e^2 V_{ud}}{2\sin^2\theta_W}\, P_W(s_{34})\, \vec{C}_{q(-;-)}(\ldots)\,,\\
\vec{C}_{(W+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) &= \frac{e^2 V_{ud}^\dagger}{2\sin^2\theta_W}\, P_W(s_{34})\, \vec{C}_{q(-;-)}(\ldots)\,.
\end{aligned}\tag{5.145}
$$

In all cases, the momentum arguments on the right-hand side are the same as on the left-hand side. Note that the $\vec{C}_{q(-;-)}$ coefficient is the same in all cases. The Wilson coefficients

270

are given by

$$\vec{C}_{x(+;+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) = A_{x,\text{fin}}(1_q^+, 2_{\bar{q}}^-; 3_\ell^+, 4_{\bar{\ell}}^-),$$

$$\vec{C}_{x(+;-)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) = \vec{C}_{x(+;+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_4, \tilde{p}_3),$$

$$\vec{C}_{q,v(-;\pm)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) = \vec{C}_{q,v(+;\pm)}(\tilde{p}_2, \tilde{p}_1; \tilde{p}_3, \tilde{p}_4),$$

$$\vec{C}_{a(-;\pm)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) = -\vec{C}_{a(+;\pm)}(\tilde{p}_2, \tilde{p}_1; \tilde{p}_3, \tilde{p}_4), \tag{5.146}$$

where $x = q, v, a$ and as discussed in Sec. 5.4 the subscript "fin" denotes the IR-finite part of the helicity amplitudes. The second relation follows from Eq. (5.138). The last two relations follow from charge conjugation invariance. At tree level and one loop only $\vec{C}_q$ receives a nonvanishing contribution. The A_q amplitude is given in Appendix B.3.1.

5.6.2 $V + 1$ Jet

$gq\bar{q}V$

For $\gamma/Z + 1$ jet, the partonic process is $gq\bar{q}\ell\bar{\ell}$, and the basis of helicity operators is

$$O_{+(+;\pm)}^{a\,\bar{\alpha}\beta} = \mathcal{B}_{1+}^a \, J_{q\,23+}^{\bar{\alpha}\beta} \, J_{\ell\,45\pm},$$

$$O_{+(-;\pm)}^{a\,\bar{\alpha}\beta} = \mathcal{B}_{1+}^a \, J_{q\,23-}^{\bar{\alpha}\beta} \, J_{\ell\,45\pm},$$

$$O_{-(+;\pm)}^{a\,\bar{\alpha}\beta} = \mathcal{B}_{1-}^a \, J_{q\,23+}^{\bar{\alpha}\beta} \, J_{\ell\,45\pm},$$

$$O_{-(-;\pm)}^{a\,\bar{\alpha}\beta} = \mathcal{B}_{1-}^a \, J_{q\,23-}^{\bar{\alpha}\beta} \, J_{\ell\,45\pm}. \tag{5.147}$$

For $W^{\mp} + 1$ jet, the partonic processes are $gu\bar{d}\ell\bar{\nu}$ and $gd\bar{u}\nu\bar{\ell}$, respectively, and the helicity operator basis is

$$O_{\pm(W^-)}^{a\,\bar{\alpha}\beta} = \mathcal{B}_{1\pm}^a \, J_{\bar{u}d\,23-}^{\bar{\alpha}\beta} \, J_{\bar{\ell}\nu\,45-},$$

$$O_{\pm(W^+)}^{a\,\bar{\alpha}\beta} = \mathcal{B}_{1\pm}^a \, J_{\bar{d}u\,23-}^{\bar{\alpha}\beta} \, J_{\bar{\nu}\ell\,45-}. \tag{5.148}$$

271

The unique color structure for $gq\bar{q}V$ is

$$\bar{T}^{a\,\alpha\bar{\beta}} = \left(T^a_{\alpha\bar{\beta}}\right),$$ (5.149)

and extracting it from each of the amplitudes, we have

$$\mathcal{I}_x(g_1 q_2 \bar{q}_3 \ell_4 \bar{\ell}_5) = \mathrm{i}\, T^{a_1}_{\alpha_2 \bar{\alpha}_3}\, A_x(1; 2_q, 3_{\bar{q}}; 4_\ell, 5_{\bar{\ell}}),$$ (5.150)

where the subscript x stands for one of q, v, a.

As for $V + 0$ jets, we write the Wilson coefficients using the electroweak decomposition in Eqs. (5.136) and (5.137). For $\gamma/Z + 1$ jet, we have

$$
\begin{aligned}
\vec{C}_{\lambda(\lambda_q;\lambda_\ell)}&(\tilde{p}_1; \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5) \\
&= e^2 \left\{ \left[Q^\ell Q^q + v^\ell_{\lambda_\ell} v^q_{\lambda_q} P_Z(s_{45}) \right] \vec{C}_{q\lambda(\lambda_q;\lambda_\ell)}(\dots) \right. \\
&\quad + \sum_{i=1}^{n_f} \left[Q^\ell Q^i + v^\ell_{\lambda_\ell} \frac{v^i_L + v^i_R}{2} P_Z(s_{45}) \right] \vec{C}_{v\lambda(\lambda_q;\lambda_\ell)}(\dots) \\
&\quad \left. + \frac{v^\ell_{\lambda_\ell}}{\sin(2\theta_W)} P_Z(s_{45}) \vec{C}_{a\lambda(\lambda_q;\lambda_\ell)}(\dots) \right\},
\end{aligned}
$$ (5.151)

where the weak couplings are determined by the helicity labels of the quark and lepton currents,

$$v^\ell_+ = v^\ell_R, \quad v^\ell_- = v^\ell_L, \qquad v^q_+ = v^q_R, \quad v^q_- = v^q_L.$$ (5.152)

For $W + 1$ jet, we have

$$\vec{C}_{\lambda(W^\mp)}(\dots) = \frac{e^2 V^{(\dagger)}_{ud}}{2\sin^2\theta_W}\, P_W(s_{45})\, \vec{C}_{q\lambda(-;-)}(\dots).$$ (5.153)

The Wilson coefficients are given by

$$\vec{C}_{x+(+;+)}(\tilde{p}_1; \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5) = A_{x,\mathrm{fin}}(1^+; 2_q^+, 3_{\bar{q}}^-; 4_\ell^+, 5_{\bar{\ell}}^-),$$

$$\vec{C}_{x\lambda(+;-)}(\tilde{p}_1; \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5) = \vec{C}_{x\lambda(+;+)}(\tilde{p}_1; \tilde{p}_2, \tilde{p}_3; \tilde{p}_5, \tilde{p}_4) ,$$

$$\vec{C}_{q,v\lambda(-;\pm)}(\tilde{p}_1; \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5) = -\vec{C}_{q,v\lambda(+;\pm)}(\tilde{p}_1; \tilde{p}_3, \tilde{p}_2; \tilde{p}_4, \tilde{p}_5) ,$$

$$\vec{C}_{a\lambda(-;\pm)}(\tilde{p}_1; \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5) = \vec{C}_{a\lambda(+;\pm)}(\tilde{p}_1; \tilde{p}_3, \tilde{p}_2; \tilde{p}_4, \tilde{p}_5) . \tag{5.154}$$

The second relation follows from Eq. (5.138), and the last two relations follow from charge conjugation invariance. The Wilson coefficients with a negative helicity gluon follow from parity invariance,

$$\vec{C}_{q,v-(+;\pm)}(\tilde{p}_1; \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5)$$
$$= \left. \vec{C}_{q,v+(-;\mp)}(\tilde{p}_1; \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5) \right|_{\langle..\rangle \leftrightarrow [..]} ,$$
$$\vec{C}_{a-(+;\pm)}(\tilde{p}_1; \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5)$$
$$= \left. -\vec{C}_{a+(-;\mp)}(\tilde{p}_1; \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5) \right|_{\langle..\rangle \leftrightarrow [..]} . \tag{5.155}$$

The helicity amplitudes for $gq\bar{q}\ell\bar{\ell}$ were calculated in Ref. [232, 233, 362]. We provide the tree-level and one-loop results in Appendix B.3.2. The two-loop amplitudes were computed in Refs. [363, 364].

$gggV$

The partonic process $ggg\ell\bar{\ell}$ first appears at one loop, and thus contributes only at relative $\mathcal{O}(\alpha_s^2)$ to $\gamma/Z + 1$ jet. Nevertheless, for the sake of completeness (and curiosity) we briefly discuss it here. The helicity operator basis is

$$O^{abc}_{+++(\pm)} = \frac{1}{3!} \mathcal{B}^a_{1+} \mathcal{B}^b_{2+} \mathcal{B}^c_{3+} J_{\ell 45\pm} ,$$

$$O^{abc}_{++-(\pm)} = \frac{1}{2} \mathcal{B}^a_{1+} \mathcal{B}^b_{2+} \mathcal{B}^c_{3-} J_{\ell 45\pm} ,$$

$$O^{abc}_{+--(\pm)} = \frac{1}{2} \mathcal{B}^a_{1+} \mathcal{B}^b_{2-} \mathcal{B}^c_{3-} J_{\ell 45\pm} ,$$

$$O^{abc}_{---(\pm)} = \frac{1}{3!} \mathcal{B}^a_{1-} \mathcal{B}^b_{2-} \mathcal{B}^c_{3-} J_{\ell 45\pm} . \tag{5.156}$$

The color space is two dimensional. We use the basis

$$\bar{T}^{abc} = \left(\mathrm{i} f^{abc} , \, d^{abc} \right) , \tag{5.157}$$

in terms of which we can write the $ggg\ell\bar{\ell}$ amplitudes as

$$\mathcal{I}_v(g_1 g_2 g_3 \ell_4 \bar{\ell}_5) = \mathrm{i}\, d^{a_1 a_2 a_3} A_v(1,2,3;4_\ell,5_{\bar{\ell}}) ,$$

$$\mathcal{I}_a(g_1 g_2 g_3 \ell_4 \bar{\ell}_5) = \mathrm{i}\, (\mathrm{i} f^{a_1 a_2 a_3}) A_a(1,2,3;4_\ell,5_{\bar{\ell}}) . \tag{5.158}$$

We will justify shortly that to all orders, only a single color structure appears for each of $\mathcal{I}_v$, $\mathcal{I}_a$. This process can only occur via a closed quark loop, so there is no $\mathcal{I}_q$ contribution. The $gggV$ operators transform under charge conjugation as

$$\mathrm{C}\, O^{abc}_{\lambda_1 \lambda_2 \lambda_3 (\pm)}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5)\, \bar{T}^{abc}\, \mathrm{C}$$

$$= O^{cba}_{\lambda_1 \lambda_2 \lambda_3 (\mp)}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3; \tilde{p}_5, \tilde{p}_4)\, \bar{T}^{abc} . \tag{5.159}$$

Charge conjugation invariance of QCD thus leads to

$$C^{abc}_{v\lambda_1 \lambda_2 \lambda_3 (\pm)}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5)$$

$$= C^{cba}_{v\lambda_1 \lambda_2 \lambda_3 (\mp)}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3; \tilde{p}_5, \tilde{p}_4)$$

$$= C^{cba}_{v\lambda_1 \lambda_2 \lambda_3 (\pm)}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5) , \tag{5.160}$$

where we used Eq. (5.138) in the last line. This implies that to all orders in the strong coupling, only the fully symmetric color structure d^{abc} can contribute to $\mathcal{I}_v$ and $\vec{C}_v$. For $\vec{C}_a$ the same relation holds but with an additional minus sign on the right-hand side due to the weak axial-vector coupling in $\mathcal{I}_a$. This implies that for $\mathcal{I}_a$ and $\vec{C}_a$ only the fully antisymmetric color structure $\mathrm{i} f^{abc}$ contributes, as given in Eq. (5.158).

We decompose the $ggg\ell\bar{\ell}$ Wilson coefficients as

$$\vec{C}_{\lambda_1\lambda_2\lambda_3(\lambda_\ell)}$$

$$= e^2 \left\{ \sum_{i=1}^{n_f} \left[Q^\ell Q^i + v^\ell_{\lambda_\ell} \frac{v^i_L + v^i_R}{2} P_Z(s_{45}) \right] \vec{C}_{v\lambda_1\lambda_2\lambda_3(\lambda_\ell)} \right.$$

$$\left. + \frac{v^\ell_{\lambda_\ell}}{\sin(2\theta_W)} P_Z(s_{45}) \, \vec{C}_{a\lambda_1\lambda_2\lambda_3(\lambda_\ell)} \right\} , \qquad (5.161)$$

where

$$v^\ell_+ = v^\ell_R , \qquad v^\ell_- = v^\ell_L , \qquad (5.162)$$

and we have

$$\vec{C}_{v\lambda_1\lambda_2\lambda_3(+)}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5)$$

$$= \begin{pmatrix} 0 \\ A_{v,\text{fin}}(1^{\lambda_1}, 2^{\lambda_2}, 3^{\lambda_3}; 4^+_\ell, 5^-_{\bar{\ell}}) \end{pmatrix} ,$$

$$\vec{C}_{a\lambda_1\lambda_2\lambda_3(+)}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5)$$

$$= \begin{pmatrix} A_{a,\text{fin}}(1^{\lambda_1}, 2^{\lambda_2}, 3^{\lambda_3}; 4^+_\ell, 5^-_{\bar{\ell}}) \\ 0 \end{pmatrix} ,$$

$$\vec{C}_{v,a\lambda_1\lambda_2\lambda_3(-)}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5)$$

$$= \vec{C}_{v,a\lambda_1\lambda_2\lambda_3(+)}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3; \tilde{p}_5, \tilde{p}_4) . \qquad (5.163)$$

For brevity, we have not written out the various gluon helicity combinations. The one-loop amplitudes for $gggZ$ were calculated in Ref. [365], and the two-loop amplitudes were computed in Ref. [366]. Since their contribution is very small we do not repeat them here.

5.6.3 $V + 2$ Jets

Here we consider the processes $q'\bar{q}'q\bar{q}\,V$, $q\bar{q}\,q\bar{q}\,V$, and $gg\,q\bar{q}\,V$. The $ggggV$ process is allowed as well, but only arises at one loop, so we do not explicitly consider here. It can be treated similarly to $gggV$, but using the $gggg$ color basis analogous to that for $ggggH$ given in Eq. (5.130).

The NLO helicity amplitudes for $V + 2$ jets were calculated in Refs. [367, 235] assuming that all kinematic scales are smaller than the top mass m_t and including the $1/m_t^2$ corrections. We give the full expressions for the LO results in Appendix B.3.3. Since the NLO results are rather long, we do not repeat them, but we show how to convert the results of Refs. [367, 235] to our notation.

$q'\bar{q}'q\bar{q}\,V$ **and** $q\bar{q}\,q\bar{q}\,V$

For $q'\bar{q}'q\bar{q}\,\ell\bar{\ell}$, the helicity operator basis is

$$O^{\bar{\alpha}\beta\bar{\gamma}\delta}_{(+;+;\pm)} = J^{\bar{\alpha}\beta}_{q'\,12+} \, J^{\bar{\gamma}\delta}_{q\,34+} \, J_{\ell\,56\pm} \,,$$

$$O^{\bar{\alpha}\beta\bar{\gamma}\delta}_{(+;-;\pm)} = J^{\bar{\alpha}\beta}_{q'\,12+} \, J^{\bar{\gamma}\delta}_{q\,34-} \, J_{\ell\,56\pm} \,,$$

$$O^{\bar{\alpha}\beta\bar{\gamma}\delta.}_{(-;+;\pm)} = J^{\bar{\alpha}\beta}_{q'\,12-} \, J^{\bar{\gamma}\delta}_{q\,34+} \, J_{\ell\,56\pm} \,,$$

$$O^{\bar{\alpha}\beta\bar{\gamma}\delta}_{(-;-;\pm)} = J^{\bar{\alpha}\beta}_{q'\,12-} \, J^{\bar{\gamma}\delta}_{q\,34-} \, J_{\ell\,56\pm} \,. \tag{5.164}$$

For identical quark flavors, $q\bar{q}\,q\bar{q}\,\ell\bar{\ell}$, the basis reduces to

$$O^{\bar{\alpha}\beta\bar{\gamma}\delta}_{(++;\pm)} = \frac{1}{4} J^{\bar{\alpha}\beta}_{q\,12+} \, J^{\bar{\gamma}\delta}_{q\,34+} \, J_{\ell\,56\pm} \,,$$

$$O^{\bar{\alpha}\beta\bar{\gamma}\delta}_{(+-;\pm)} = J^{\bar{\alpha}\beta}_{q\,12+} \, J^{\bar{\gamma}\delta}_{q\,34-} \, J_{\ell\,56\pm} \,,$$

$$O^{\bar{\alpha}\beta\bar{\gamma}\delta}_{(--;\pm)} = \frac{1}{4} J^{\bar{\alpha}\beta}_{q\,12-} \, J^{\bar{\gamma}\delta}_{q\,34-} \, J_{\ell\,56\pm} \,. \tag{5.165}$$

276

For $W + 2$ jets, the corresponding partonic processes are $q\bar{q}\,ud\,\ell\bar{\nu}$ and $q\bar{q}\,d\bar{u}\,\nu\bar{\ell}$, and the helicity operator basis is

$$O^{\bar{\alpha}\beta\bar{\gamma}\delta}_{(\pm;W-)} = J^{\bar{\alpha}\beta}_{q\,12\pm}\,J^{\bar{\gamma}\delta}_{\bar{u}d\,34-}\,J_{\ell\nu\,56-}\,,$$

$$O^{\bar{\alpha}\beta\bar{\gamma}\delta}_{(\pm;W+)} = J^{\bar{\alpha}\beta}_{q\,12\pm}\,J^{\bar{\gamma}\delta}_{\bar{d}u\,34-}\,J_{\bar{\nu}\ell\,56-}\,. \tag{5.166}$$

We use the color basis

$$\bar{T}^{\alpha\bar{\beta}\gamma\bar{\delta}} = 2T_F\Big(\delta_{\alpha\bar{\delta}}\,\delta_{\gamma\bar{\beta}}\,,\ \delta_{\alpha\bar{\beta}}\,\delta_{\gamma\bar{\delta}}\Big)\,. \tag{5.167}$$

For distinct quark flavors, the color decomposition of the amplitudes in this basis is

$$\mathcal{I}_x(q_1'\bar{q}_2'q_3\bar{q}_4\ell_5\bar{\ell}_6) \tag{5.168}$$

$$= 2T_F\mathrm{i}\,\delta_{\alpha_1\bar{\alpha}_4}\delta_{\alpha_3\bar{\alpha}_2}\,A_x(1_{q'},2_{\bar{q}'};3_q,4_{\bar{q}};5_\ell,6_{\bar{\ell}})$$

$$+ 2T_F\mathrm{i}\,\delta_{\alpha_1\bar{\alpha}_2}\delta_{\alpha_3\bar{\alpha}_4}\,\frac{1}{N}\,B_x(1_{q'},2_{\bar{q}'};3_q,4_{\bar{q}};5_\ell,6_{\bar{\ell}})\,.$$

For identical quark flavors the amplitudes can be obtained from the distinct flavor amplitudes using

$$\mathcal{I}_x(q_1\bar{q}_2q_3\bar{q}_4\ell_5\bar{\ell}_6) = \mathcal{I}_x(q_1'\bar{q}_2'q_3\bar{q}_4\ell_5\bar{\ell}_6)$$

$$- \mathcal{I}_x(q_1'\bar{q}_4'q_3\bar{q}_2\ell_5\bar{\ell}_6)\,, \tag{5.169}$$

where it is to be understood that the electroweak couplings of q' must also be replaced by those of q.

Writing the Wilson coefficients in the decomposition in Eqs. (5.136) and (5.137), we have for the $q'\bar{q}'q\bar{q}\,\ell\bar{\ell}$ channel

$$\vec{C}_{(\lambda_{q'};\lambda_q;\lambda_\ell)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_5,\tilde{p}_6)$$

$$= e^2\left\{\big[Q^\ell Q^q + v^\ell_{\lambda_\ell}v^q_{\lambda_q}P_Z(s_{56})\big]\right.$$

$$\times \vec{C}_{q(\lambda_{q'};\lambda_q;\lambda_\ell)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_5,\tilde{p}_6)$$

277

$$+ \left[Q^\ell Q^{q'} + v^\ell_{\lambda_\ell} v^{q'}_{\lambda_{q'}} P_Z(s_{56}) \right]$$

$$\times \, \vec{C}_{q(\lambda_q ; \lambda_{q'} ; \lambda_\ell)}(\tilde{p}_3, \tilde{p}_4; \tilde{p}_1, \tilde{p}_2; \tilde{p}_5, \tilde{p}_6)$$

$$+ \sum_{i=1}^{n_f} \left[Q^\ell Q^i + v^\ell_{\lambda_\ell} \frac{v^i_L + v^i_R}{2} P_Z(s_{56}) \right] \vec{C}_{v(\lambda_{q'} ; \lambda_q ; \lambda_\ell)}(\dots)$$

$$+ \left. \frac{v^\ell_{\lambda_\ell}}{\sin(2\theta_W)} P_Z(s_{56}) \, \vec{C}_{a(\lambda_{q'} ; \lambda_q ; \lambda_\ell)}(\dots) \right\}, \tag{5.170}$$

with the weak couplings

$$v^\ell_+ = v^\ell_R, \quad v^\ell_- = v^\ell_L, \qquad v^q_+ = v^q_R, \quad v^q_- = v^q_L. \tag{5.171}$$

The same decomposition is used for the case of identical flavors, $q\bar{q}\,q\bar{q}\,\ell\bar{\ell}$. For the $W^\mp$ channels, $q\bar{q}\,u\bar{d}\,\ell\bar{\nu}$ and $q\bar{q}\,d\bar{u}\,\nu\bar{\ell}$, we have

$$\vec{C}_{(\lambda_q ; W^\mp)}(\dots) = \frac{e^2 V^{(\dagger)}_{ud}}{2 \sin^2 \theta_W} \, P_W(s_{56}) \, \vec{C}_{q(\lambda_q ; - ; -)}(\dots). \tag{5.172}$$

The coefficients for $q'\bar{q}'q\bar{q}V$ are given by

$$\vec{C}_{x(+;+;+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5, \tilde{p}_6) = \begin{pmatrix} A_{x,\mathrm{fin}}(1^+_{q'}, 2^-_{\bar{q}'}; 3^+_q, 4^-_{\bar{q}}; 5^+_\ell, 6^-_{\bar{\ell}}) \\ \frac{1}{N} B_{x,\mathrm{fin}}(1^+_{q'}, 2^-_{\bar{q}'}; 3^+_q, 4^-_{\bar{q}}; 5^+_\ell, 6^-_{\bar{\ell}}) \end{pmatrix},$$

$$\vec{C}_{x(+;-;+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5, \tilde{p}_6) = \begin{pmatrix} A_{x,\mathrm{fin}}(1^+_{q'}, 2^-_{\bar{q}'}; 3^-_q, 4^+_{\bar{q}}; 5^+_\ell, 6^-_{\bar{\ell}}) \\ \frac{1}{N} B_{x,\mathrm{fin}}(1^+_{q'}, 2^-_{\bar{q}'}; 3^-_q, 4^+_{\bar{q}}; 5^+_\ell, 6^-_{\bar{\ell}}) \end{pmatrix},$$

$$\vec{C}_{x(+;\pm;-)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5, \tilde{p}_6) = \vec{C}_{x(+;\pm;+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_6, \tilde{p}_5),$$

$$\vec{C}_{q,v(-;+;\pm)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5, \tilde{p}_6) = -\vec{C}_{q,v(+;-;\pm)}(\tilde{p}_2, \tilde{p}_1; \tilde{p}_4, \tilde{p}_3; \tilde{p}_5, \tilde{p}_6),$$

$$\vec{C}_{a(-;+;\pm)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5, \tilde{p}_6) = \vec{C}_{a(+;-;\pm)}(\tilde{p}_2, \tilde{p}_1; \tilde{p}_4, \tilde{p}_3; \tilde{p}_5, \tilde{p}_6),$$

$$\vec{C}_{q,v(-;-;\pm)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5, \tilde{p}_6) = -\vec{C}_{q,v(+;+;\pm)}(\tilde{p}_2, \tilde{p}_1; \tilde{p}_4, \tilde{p}_3; \tilde{p}_5, \tilde{p}_6),$$

$$\vec{C}_{a(-;-;\pm)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4; \tilde{p}_5, \tilde{p}_6) = \vec{C}_{a(+;+;\pm)}(\tilde{p}_2, \tilde{p}_1; \tilde{p}_4, \tilde{p}_3; \tilde{p}_5, \tilde{p}_6), \tag{5.173}$$

and for $q\bar{q}\,q\bar{q}\,V$ they are given in terms of the amplitudes $A_{x,\text{fin}}$ and $B_{x,\text{fin}}$ for $q'\bar{q}'q\bar{q}\,V$ by

$$\vec{C}_{x(++;+)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_5,\tilde{p}_6) = \begin{pmatrix} A_{x,\text{fin}}(1_q^+,2_{\bar{q}}^-;3_q^+,4_{\bar{q}}^-;5_\ell^+,6_{\bar{\ell}}^-) - \frac{1}{N}B_{x,\text{fin}}(1_q^+,4_{\bar{q}}^-;3_q^+,2_{\bar{q}}^-;5_\ell^+,6_{\bar{\ell}}^-) \\ \frac{1}{N}B_{x,\text{fin}}(1_q^+,2_{\bar{q}}^-;3_q^+,4_{\bar{q}}^-;5_\ell^+,6_{\bar{\ell}}^-) - A_{x,\text{fin}}(1_q^+,4_{\bar{q}}^-;3_q^+,2_{\bar{q}}^-;5_\ell^+,6_{\bar{\ell}}^-) \end{pmatrix},$$

$$\vec{C}_{x(+-;+)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_5,\tilde{p}_6) = \begin{pmatrix} A_{x,\text{fin}}(1_q^+,2_{\bar{q}}^-;3_q^-,4_{\bar{q}}^+;5_\ell^+,6_{\bar{\ell}}^-) \\ \frac{1}{N}B_{x,\text{fin}}(1_q^+,2_{\bar{q}}^-;3_q^-,4_{\bar{q}}^+;5_\ell^+,6_{\bar{\ell}}^-) \end{pmatrix},$$

$$\vec{C}_{x(+\pm;-)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_5,\tilde{p}_6) = \vec{C}_{x(+\pm;+)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_6,\tilde{p}_5)\,,$$

$$\vec{C}_{q,v(--;\pm)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_5,\tilde{p}_6) = -\vec{C}_{q,v(++;\pm)}(\tilde{p}_2,\tilde{p}_1;\tilde{p}_4,\tilde{p}_3;\tilde{p}_5,\tilde{p}_6)\,,$$

$$\vec{C}_{a(--;\pm)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_5,\tilde{p}_6) = \vec{C}_{a(++;\pm)}(\tilde{p}_2,\tilde{p}_1;\tilde{p}_4,\tilde{p}_3;\tilde{p}_5,\tilde{p}_6)\,. \tag{5.174}$$

The various relations for the coefficients with flipped helicities follow from Eq. (5.138) and charge conjugation invariance. The tree-level helicity amplitudes are given in Appendix B.3.3.

$gg\,q\bar{q}\,V$

For $gg\,q\bar{q}\,\ell\bar{\ell}$, the helicity operator basis consists of 12 independent operators,

$$O^{ab\,\bar{\alpha}\beta}_{++(+;\pm)} = \frac{1}{2}\,\mathcal{B}^a_{1+}\,\mathcal{B}^b_{2+}\,J^{\bar{\alpha}\beta}_{q\,34+}\,J_{\ell\,56\pm}\,,$$

$$O^{ab\,\bar{\alpha}\beta}_{++(-;\pm)} = \frac{1}{2}\,\mathcal{B}^a_{1+}\,\mathcal{B}^b_{2+}\,J^{\bar{\alpha}\beta}_{q\,34-}\,J_{\ell\,56\pm}\,,$$

$$O^{ab\,\bar{\alpha}\beta}_{+-(+;\pm)} = \mathcal{B}^a_{1+}\,\mathcal{B}^b_{2-}\,J^{\bar{\alpha}\beta}_{q\,34+}\,J_{\ell\,56\pm}\,,$$

$$O^{ab\,\bar{\alpha}\beta}_{+-(-;\pm)} = \mathcal{B}^a_{1+}\,\mathcal{B}^b_{2-}\,J^{\bar{\alpha}\beta}_{q\,34-}\,J_{\ell\,56\pm}\,,$$

$$O^{ab\,\bar{\alpha}\beta}_{--(+;\pm)} = \frac{1}{2}\,\mathcal{B}^a_{1-}\,\mathcal{B}^b_{2-}\,J^{\bar{\alpha}\beta}_{q\,34+}\,J_{\ell\,56\pm}\,,$$

$$O^{ab\,\bar{\alpha}\beta}_{--(-;\pm)} = \frac{1}{2}\,\mathcal{B}^a_{1-}\,\mathcal{B}^b_{2-}\,J^{\bar{\alpha}\beta}_{q\,34-}\,J_{\ell\,56\pm}\,. \tag{5.175}$$

279

For $W^{\mp}$, the corresponding partonic processes are $gg\,u\bar{d}\,\ell\bar{\nu}$ and $gg\,d\bar{u}\,\nu\bar{\ell}$, and the helicity operator basis reduces to six independent operators,

$$O^{ab\,\bar{\alpha}\beta}_{++(W-)} = \frac{1}{2}\,\mathcal{B}^a_{1+}\,\mathcal{B}^b_{2+}\,J^{\bar{\alpha}\beta}_{\bar{u}d\,34-}\,J_{\bar{\ell}\nu\,56-}\,,$$

$$O^{ab\,\bar{\alpha}\beta}_{+-(W-)} = \mathcal{B}^a_{1+}\,\mathcal{B}^b_{2-}\,J^{\bar{\alpha}\beta}_{\bar{u}d\,34-}\,J_{\bar{\ell}\nu\,56-}\,,$$

$$O^{ab\,\bar{\alpha}\beta}_{--(W-)} = \frac{1}{2}\,\mathcal{B}^a_{1-}\,\mathcal{B}^b_{2-}\,J^{\bar{\alpha}\beta}_{\bar{u}d\,34-}\,J_{\bar{\ell}\nu\,56-}\,,$$

$$O^{ab\,\bar{\alpha}\beta}_{++(W+)} = \frac{1}{2}\,\mathcal{B}^a_{1+}\,\mathcal{B}^b_{2+}\,J^{\bar{\alpha}\beta}_{\bar{d}u\,34-}\,J_{\bar{\nu}\ell\,56-}\,,$$

$$O^{ab\,\bar{\alpha}\beta}_{+-(W+)} = \mathcal{B}^a_{1+}\,\mathcal{B}^b_{2-}\,J^{\bar{\alpha}\beta}_{\bar{d}u\,34-}\,J_{\bar{\nu}\ell\,56-}\,,$$

$$O^{ab\,\bar{\alpha}\beta}_{--(W+)} = \frac{1}{2}\,\mathcal{B}^a_{1-}\,\mathcal{B}^b_{2-}\,J^{\bar{\alpha}\beta}_{\bar{d}u\,34-}\,J_{\bar{\nu}\ell\,56-}\,. \tag{5.176}$$

We use the color basis

$$\bar{T}^{ab\alpha\bar{\beta}} = \left((T^aT^b)_{\alpha\bar{\beta}}\,,\,(T^bT^a)_{\alpha\bar{\beta}}\,,\,\mathrm{tr}[T^aT^b]\,\delta_{\alpha\bar{\beta}}\right), \tag{5.177}$$

and the amplitudes are color-decomposed as

$$\mathcal{I}_x(g_1g_2q_3\bar{q}_4\ell_5\bar{\ell}_6)$$
$$= \mathrm{i}\sum_{\sigma\in S_2}\left[T^{a_{\sigma(1)}}T^{a_{\sigma(2)}}\right]_{\alpha_3\bar{\alpha}_4}A_x(\sigma(1),\sigma(2);3_q,4_{\bar{q}};5_\ell,6_{\bar{\ell}})$$
$$+ \mathrm{i}\,\mathrm{tr}[T^{a_1}T^{a_2}]\,\delta_{\alpha_3\bar{\alpha}_4}\,B_x(1,2;3_q,4_{\bar{q}};5_\ell,6_{\bar{\ell}})\,. \tag{5.178}$$

Writing the Wilson coefficients in the decomposition in Eqs. (5.136) and (5.137), we have for the $gg\,q\bar{q}\,\ell\bar{\ell}$ channel

$$\vec{C}_{\lambda_1\lambda_2(\lambda_q;\lambda_\ell)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_5,\tilde{p}_6)$$
$$= e^2\left\{\left[Q^\ell Q^q + v^\ell_{\lambda_\ell}v^q_{\lambda_q}P_Z(s_{56})\right]\vec{C}_{q\lambda_1\lambda_2(\lambda_q;\lambda_\ell)}(\ldots)\right.$$
$$\left.+ \sum_{i=1}^{n_f}\left[Q^\ell Q^i + v^\ell_{\lambda_\ell}\frac{v^i_L + v^i_R}{2}P_Z(s_{56})\right]\vec{C}_{v\lambda_1\lambda_2(\lambda_q;\lambda_\ell)}(\ldots)\right.$$

$$+ \frac{v^\ell_{\lambda_\ell}}{\sin(2\theta_W)} P_Z(s_{56})\, \vec{C}_{a\lambda_1\lambda_2(\lambda_q;\lambda_\ell)}(\dots) \bigg\} \,, \tag{5.179}$$

with the weak couplings

$$v^\ell_+ = v^\ell_R \,, \quad v^\ell_- = v^\ell_L \,, \qquad v^q_+ = v^q_R \,, \quad v^q_- = v^q_L \,. \tag{5.180}$$

For the $W^\mp$ channels $gg\,u\bar{d}\,\ell\bar{\nu}$ and $gg\,d\bar{u}\,\nu\bar{\ell}$, we have

$$\vec{C}_{\lambda_1\lambda_2(W^\mp)}(\dots) = \frac{e^2 V^{(\dagger)}_{ud}}{2\sin^2\theta_W}\, P_W(s_{56})\, \vec{C}_{q\lambda_1\lambda_2(-;-)}(\dots)\,. \tag{5.181}$$

The coefficients for $gg\,q\bar{q}\,V$ are then given by

$$\vec{C}_{x\lambda_1\lambda_2(+;+)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_5,\tilde{p}_6)$$

$$= \begin{pmatrix} A_{x,\text{fin}}(1^{\lambda_1}, 2^{\lambda_2}; 3^+_q, 4^-_{\bar{q}}; 5^+_\ell, 6^-_{\bar{\ell}}) \\ A_{x,\text{fin}}(2^{\lambda_2}, 1^{\lambda_1}; 3^+_q, 4^-_{\bar{q}}; 5^+_\ell, 6^-_{\bar{\ell}}) \\ B_{x,\text{fin}}(1^{\lambda_1}, 2^{\lambda_2}; 3^+_q, 4^-_{\bar{q}}; 5^+_\ell, 6^-_{\bar{\ell}}) \end{pmatrix} \,,$$

$$\vec{C}_{x\lambda_1\lambda_2(+;-)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_5,\tilde{p}_6)$$

$$= \vec{C}_{x\lambda_1\lambda_2(+;+)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_6,\tilde{p}_5)\,. \tag{5.182}$$

The remaining Wilson coefficients are obtained by charge conjugation invariance as follows,

$$\vec{C}_{q,v\lambda_1\lambda_2(-;\pm)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_5,\tilde{p}_6)$$

$$= \widehat{V}\vec{C}_{q,v\lambda_1\lambda_2(+;\pm)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_4,\tilde{p}_3;\tilde{p}_5,\tilde{p}_6)\,,$$

$$\vec{C}_{a\lambda_1\lambda_2(-;\pm)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4;\tilde{p}_5,\tilde{p}_6)$$

$$= -\widehat{V}\vec{C}_{a\lambda_1\lambda_2(+;\pm)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_4,\tilde{p}_3;\tilde{p}_5,\tilde{p}_6)\,,$$

$$\text{with} \quad \widehat{V} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\,. \tag{5.183}$$

The tree-level helicity amplitudes are given in Appendix B.3.3.

5.7 $pp \to$ Jets

In this section, we give the operator basis and matching relations for $pp \to 2, 3$ jets. We consider only the QCD contributions, so that quarks only appear in same-flavor quark-antiquark pairs with the same chirality, and so are described by the currents $J_{ij\pm}$. The helicity amplitudes for each channel are given in Appendix B.4.

5.7.1 $pp \to$ 2 Jets

For $pp \to 2$ jets, the partonic channels $q\bar{q}\,q'\bar{q}'$, $q\bar{q}\,q\bar{q}$, $q\bar{q}gg$, and $gggg$ contribute. We will discuss each in turn. The one-loop helicity amplitudes for all partonic channels were first calculated in Ref. [244]. The tree-level and one-loop results are given in Appendix B.4.1. The two-loop amplitudes have also been calculated, and can be found in Refs. [368, 369] for $q\bar{q}gg$, Refs. [370, 371, 372, 373] for $q\bar{q}\,q'\bar{q}'$, $q\bar{q}\,q\bar{q}$ and in Refs. [374, 375] for $gggg$.

$q\bar{q}\,q'\bar{q}'$ and $q\bar{q}\,q\bar{q}$

In the case of distinct quark flavors, $q\bar{q}\,q'\bar{q}'$, the helicity basis consists of four independent operators,

$$
\begin{aligned}
O_{(+;+)}^{\bar{\alpha}\beta\bar{\gamma}\delta} &= J_{q\,12+}^{\bar{\alpha}\beta}\, J_{q'\,34+}^{\bar{\gamma}\delta}\,, \\
O_{(+;-)}^{\bar{\alpha}\beta\bar{\gamma}\delta} &= J_{q\,12+}^{\bar{\alpha}\beta}\, J_{q'\,34-}^{\bar{\gamma}\delta}\,, \\
O_{(-;+)}^{\bar{\alpha}\beta\bar{\gamma}\delta} &= J_{q\,12-}^{\bar{\alpha}\beta}\, J_{q'\,34+}^{\bar{\gamma}\delta}\,, \\
O_{(-;-)}^{\bar{\alpha}\beta\bar{\gamma}\delta} &= J_{q\,12-}^{\bar{\alpha}\beta}\, J_{q'\,34-}^{\bar{\gamma}\delta}\,.
\end{aligned}
\tag{5.184}
$$

For identical quark flavors, $q\bar{q}\,q\bar{q}$, the helicity basis only has three independent operators,

$$
O_{(++)}^{\bar{\alpha}\beta\bar{\gamma}\delta} = \frac{1}{4}\, J_{12+}^{\bar{\alpha}\beta}\, J_{34+}^{\bar{\gamma}\delta}\,,
$$

282

$$O^{\bar{\alpha}\beta\bar{\gamma}\delta}_{(+-)} = J^{\bar{\alpha}\beta}_{12+} \, J^{\bar{\gamma}\delta}_{34-} \,,$$

$$O^{\bar{\alpha}\beta\bar{\gamma}\delta}_{(--)} = \frac{1}{4} \, J^{\bar{\alpha}\beta}_{12-} \, J^{\bar{\gamma}\delta}_{34-} \,. \tag{5.185}$$

Here we have not made the flavor label explicit, since both quark currents have the same flavor. In both cases we use the color basis

$$\bar{T}^{\alpha\bar{\beta}\gamma\bar{\delta}} = 2T_F \left(\delta_{\alpha\bar{\delta}} \, \delta_{\gamma\bar{\beta}} \,,\; \delta_{\alpha\bar{\beta}} \, \delta_{\gamma\bar{\delta}} \right) \,. \tag{5.186}$$

The QCD helicity amplitudes for $q\bar{q}\,q'\bar{q}'$ can be color-decomposed in the basis of Eq. (5.186) as

$$\mathcal{I}(q_1\bar{q}_2 q'_3 \bar{q}'_4) = 2T_F \mathrm{i} \Big[\delta_{\alpha_1\bar{\alpha}_4}\delta_{\alpha_3\bar{\alpha}_2} A(1_q, 2_{\bar{q}}; 3_{q'}, 4_{\bar{q}'} \tag{5.187}$$

$$+ \frac{1}{N} \delta_{\alpha_1\bar{\alpha}_2}\delta_{\alpha_3\bar{\alpha}_4} B(1_q, 2_{\bar{q}}; 3_{q'}, 4_{\bar{q}'}) \Big] \,,$$

where we have included a factor of $1/N$ for convenience. The amplitude vanishes in the case that the quark and antiquark of the same flavor have the same helicity. This is equivalent to the fact that the operators of Eq. (5.184) provide a complete basis of helicity operators. For identical quark flavors, the QCD amplitudes can be written in terms of the amplitudes for the distinct flavor case as

$$\mathcal{I}(q_1\bar{q}_2 q_3 \bar{q}_4) = \mathcal{I}(q_1\bar{q}_2 q'_3 \bar{q}'_4) - \mathcal{I}(q_1\bar{q}'_4 q'_3 \bar{q}_2) \,. \tag{5.188}$$

The Wilson coefficients for $q\bar{q}\,q'\bar{q}'$ are then given by

$$\vec{C}_{(+;+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) = \begin{pmatrix} A_{\mathrm{fin}}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-) \\ \frac{1}{N} B_{\mathrm{fin}}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-) \end{pmatrix} \,,$$

$$\vec{C}_{(+;-)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) = \begin{pmatrix} A_{\mathrm{fin}}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+) \\ \frac{1}{N} B_{\mathrm{fin}}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+) \end{pmatrix} \,,$$

$$\vec{C}_{(-;+)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) = \vec{C}_{(+;-)}(\tilde{p}_2, \tilde{p}_1; \tilde{p}_4, \tilde{p}_3) \,,$$

283

$$\vec{C}_{(-;-)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) = \vec{C}_{(+;+)}(\tilde{p}_2, \tilde{p}_1; \tilde{p}_4, \tilde{p}_3)\,, \tag{5.189}$$

and for $q\bar{q}\,q\bar{q}$ they are given in terms of the amplitudes A_{fin} and B_{fin} for $q\bar{q}\,q'\bar{q}'$ by

$$\vec{C}_{(++)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) = \begin{pmatrix} A_{\text{fin}}(1_q^+, 2_{\bar{q}}^-; 3_q^+, 4_{\bar{q}}^-) - \frac{1}{N} B_{\text{fin}}(1_q^+, 4_{\bar{q}}^-; 3_q^+, 2_{\bar{q}}^-) \\ \frac{1}{N} B_{\text{fin}}(1_q^+, 2_{\bar{q}}^-; 3_q^+, 4_{\bar{q}}^-) - A_{\text{fin}}(1_q^+, 4_{\bar{q}}^-; 3_q^+, 2_{\bar{q}}^-) \end{pmatrix},$$

$$\vec{C}_{(+-)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) = \begin{pmatrix} A_{\text{fin}}(1_q^+, 2_{\bar{q}}^-; 3_q^-, 4_{\bar{q}}^+) \\ \frac{1}{N} B_{\text{fin}}(1_q^+, 2_{\bar{q}}^-; 3_q^-, 4_{\bar{q}}^+) \end{pmatrix},$$

$$\vec{C}_{(--)}(\tilde{p}_1, \tilde{p}_2; \tilde{p}_3, \tilde{p}_4) = \vec{C}_{(++)}(\tilde{p}_2, \tilde{p}_1; \tilde{p}_4, \tilde{p}_3)\,. \tag{5.190}$$

The relations for $\vec{C}_{(-;\pm)}$ and $\vec{C}_{(--)}$ follow from charge conjugation invariance. The Wilson coefficient $\vec{C}_{(+-)}$ is equal to $\vec{C}_{(+;-)}$, since the amplitude vanishes when the quark and anti-quark of the same flavor have the same helicity (both $+$ or both $-$), so there is no exchange term. The subscript "fin" in Eqs. (5.189) and (5.190) denotes the IR-finite part of the helicity amplitudes as discussed in Sec. 5.4, see Eq. (5.96). Recall that the symmetry factors of $1/4$ in Eq. (5.185) already take care of the interchange of identical (anti)quarks, so there are no additional symmetry factors needed for $\vec{C}_{(++)}$. Explicit expressions for all required partial amplitudes at tree level and one loop are given in Appendix B.4.1.

$ggq\bar{q}$

For $ggq\bar{q}$, the helicity basis has a total of six independent operators,

$$O_{++(+)}^{ab\,\bar{\alpha}\beta} = \frac{1}{2}\, \mathcal{B}_{1+}^a\, \mathcal{B}_{2+}^b\, J_{34+}^{\bar{\alpha}\beta}\,,$$

$$O_{+-(+)}^{ab\,\bar{\alpha}\beta} = \mathcal{B}_{1+}^a\, \mathcal{B}_{2-}^b\, J_{34+}^{\bar{\alpha}\beta}\,,$$

$$O_{--(+)}^{ab\,\bar{\alpha}\beta} = \frac{1}{2}\mathcal{B}_{1-}^a\, \mathcal{B}_{2-}^b\, J_{34+}^{\bar{\alpha}\beta}\,,$$

$$O_{++(-)}^{ab\,\bar{\alpha}\beta} = \frac{1}{2}\, \mathcal{B}_{1+}^a\, \mathcal{B}_{2+}^b\, J_{34-}^{\bar{\alpha}\beta}\,,$$

$$O_{+-(-)}^{ab\,\bar{\alpha}\beta} = \mathcal{B}_{1+}^a\, \mathcal{B}_{2-}^b\, J_{34-}^{\bar{\alpha}\beta}\,,$$

$$O^{ab\,\bar{\alpha}\beta}_{--(-)} = \frac{1}{2}\mathcal{B}^a_{1-}\,\mathcal{B}^b_{2-}\,J^{\bar{\alpha}\beta}_{34-}\,.\tag{5.191}$$

Note that the use of a helicity basis has made it easy to count the number of required operators. [9] For the color structure, we use the basis

$$\bar{T}^{ab\alpha\bar{\beta}} = \left((T^aT^b)_{\alpha\bar{\beta}}\,,\,(T^bT^a)_{\alpha\bar{\beta}}\,,\,\mathrm{tr}[T^aT^b]\,\delta_{\alpha\bar{\beta}}\right).\tag{5.192}$$

The color decomposition of the QCD helicity amplitudes into partial amplitudes using the color basis of Eq. (5.192) is

$$\begin{aligned}
\mathcal{I}&(g_1g_2\,q_3\bar{q}_4)\\
&= \mathrm{i}\sum_{\sigma\in S_2}\left[T^{a_{\sigma(1)}}T^{a_{\sigma(2)}}\right]_{\alpha_3\bar{\alpha}_4}A(\sigma(1),\sigma(2);3_q,4_{\bar{q}})\\
&\quad + \mathrm{i}\,\mathrm{tr}[T^{a_1}T^{a_2}]\,\delta_{\alpha_3\bar{\alpha}_4}B(1,2;3_q,4_{\bar{q}})\,,
\end{aligned}\tag{5.193}$$

from which we can read off the Wilson coefficients,

$$\vec{C}_{+-(+)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4) = \begin{pmatrix} A_{\mathrm{fin}}(1^+,2^-;3_q^+,4_{\bar{q}}^-)\\ A_{\mathrm{fin}}(2^-,1^+;3_q^+,4_{\bar{q}}^-)\\ B_{\mathrm{fin}}(1^+,2^-;3_q^+,4_{\bar{q}}^-) \end{pmatrix},$$

$$\vec{C}_{++(+)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4) = \begin{pmatrix} A_{\mathrm{fin}}(1^+,2^+;3_q^+,4_{\bar{q}}^-)\\ A_{\mathrm{fin}}(2^+,1^+;3_q^+,4_{\bar{q}}^-)\\ B_{\mathrm{fin}}(1^+,2^+;3_q^+,4_{\bar{q}}^-) \end{pmatrix},$$

$$\vec{C}_{--(+)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4) = \begin{pmatrix} A_{\mathrm{fin}}(1^-,2^-;3_q^+,4_{\bar{q}}^-)\\ A_{\mathrm{fin}}(2^-,1^-;3_q^+,4_{\bar{q}}^-)\\ B_{\mathrm{fin}}(1^-,2^-;3_q^+,4_{\bar{q}}^-) \end{pmatrix}.\tag{5.194}$$

---

[9]This should be contrasted with the more complicated basis given in Eq. (126) of Ref. [29] which is built from fields χ_{n_i} and $\mathcal{B}^{\perp\mu}_{n_i}$ and standard Dirac structures. It can be reduced to a minimal basis using identities such as $O_2 = -O_1$, $O_8 = O_7 + 4tO_3 - 4tO_4$ and $O_6 = O_5 - 2O_1 + \mathcal{O}(\epsilon)$ where $t = -\omega_1\omega_3 n_1\cdot n_3/2$, and then can be related to the basis used here.

The remaining coefficients follow from charge conjugation as discussed in Sec. 5.3.5,

$$\vec{C}_{\lambda_1\lambda_2(-)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_3,\tilde{p}_4) = \widehat{V}\vec{C}_{\lambda_1\lambda_2(+)}(\tilde{p}_1,\tilde{p}_2;\tilde{p}_4,\tilde{p}_3)\,,$$

$$\text{with}\qquad \widehat{V} = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}\,. \qquad (5.195)$$

At tree level, the partial amplitudes are well known, and only the first two entries in $\vec{C}_{+-(\pm)}$ are nonzero. Explicit expressions for all amplitudes at tree level and one loop are given in Appendix B.4.1.

gggg

For *gggg*, the helicity basis has five independent operators,

$$\begin{aligned}
O^{abcd}_{++++} &= \frac{1}{4!}\,\mathcal{B}^a_{1+}\mathcal{B}^b_{2+}\mathcal{B}^c_{3+}\mathcal{B}^d_{4+}\,, \\
O^{abcd}_{+++-} &= \frac{1}{3!}\,\mathcal{B}^a_{1+}\mathcal{B}^b_{2+}\mathcal{B}^c_{3+}\mathcal{B}^d_{4-}\,, \\
O^{abcd}_{++--} &= \frac{1}{4}\,\mathcal{B}^a_{1+}\mathcal{B}^b_{2+}\mathcal{B}^c_{3-}\mathcal{B}^d_{4-}\,, \\
O^{abcd}_{---+} &= \frac{1}{3!}\,\mathcal{B}^a_{1-}\mathcal{B}^b_{2-}\mathcal{B}^c_{3-}\mathcal{B}^d_{4+}\,, \\
O^{abcd}_{----} &= \frac{1}{4!}\,\mathcal{B}^a_{1-}\mathcal{B}^b_{2-}\mathcal{B}^c_{3-}\mathcal{B}^d_{4-}\,.
\end{aligned} \qquad (5.196)$$

We use the color basis

$$\bar{T}^{abcd} = \frac{1}{2\cdot 2T_F}\begin{pmatrix} \text{tr}[abcd] + \text{tr}[dcba] \\ \text{tr}[acdb] + \text{tr}[bdca] \\ \text{tr}[adbc] + \text{tr}[cbda] \\ 2\text{tr}[ab]\text{tr}[cd] \\ 2\text{tr}[ac]\text{tr}[db] \\ 2\text{tr}[ad]\text{tr}[bc] \end{pmatrix}^T\,, \qquad (5.197)$$

286

where we have used the shorthand notation

$$\text{tr}[ab] = \text{tr}[T^a T^b], \qquad \text{tr}[abcd] = \text{tr}[T^a T^b T^c T^d]. \tag{5.198}$$

Under charge conjugation, the operators transform as

$$C \, O^{abcd}_{\lambda_1\lambda_2\lambda_3\lambda_4} \, \bar{T}^{abcd} \, C = O^{dcba}_{\lambda_1\lambda_2\lambda_3\lambda_4} \bar{T}^{abcd}. \tag{5.199}$$

Thus, charge conjugation invariance of QCD leads to

$$C^{abcd}_{\lambda_1\lambda_2\lambda_3\lambda_4} = C^{dcba}_{\lambda_1\lambda_2\lambda_3\lambda_4}. \tag{5.200}$$

In principle, there are three more color structures with a minus sign instead of the plus sign in the first three lines in Eq. (5.197). Since charge conjugation is a symmetry of QCD, Eq. (5.200) holds to all orders, so these additional color structures cannot contribute. In particular, the color structures in Eq. (5.197) cannot mix into these additional structures at any order. Hence, it is sufficient to consider the reduced basis in Eq. (5.197) instead of the 9 different color structures, which were used for example in Ref. [296]. Note that for $N = 3$ it is possible to further reduce the color basis by one using the relation

$$\text{tr}[abcd + dcba] + \text{tr}[acdb + bdca] + \text{tr}[adbc + cbda]$$
$$= \text{tr}[ab]\text{tr}[cd] + \text{tr}[ac]\text{tr}[db] + \text{tr}[ad]\text{tr}[bc]. \tag{5.201}$$

We refrain from doing so, since it makes the structure of the anomalous dimension matrix less visible, and because there are no such relations for $N > 3$.

The color decomposition of the QCD amplitude into partial amplitudes using the color basis in Eq. (5.197) is

$$\mathcal{I}(g_1 g_2 g_3 g_4) = \frac{i}{2T_F} \left[\sum_{\sigma \in S_4/Z_4} \text{tr}[a_{\sigma(1)} a_{\sigma(2)} a_{\sigma(3)} a_{\sigma(4)}] \right.$$

$$\times\, A\big(\sigma(1),\sigma(2),\sigma(3),\sigma(4)\big)$$

$$+ \sum_{\sigma \in S_4/Z_2^3} \mathrm{tr}[a_{\sigma(1)}a_{\sigma(2)}]\mathrm{tr}[a_{\sigma(3)}a_{\sigma(4)}]$$

$$\times\, B\big(\sigma(1),\sigma(2),\sigma(3),\sigma(4)\big)\bigg],$$

$$(5.202)$$

from which we obtain the Wilson coefficients

$$\vec{C}_{++--}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3,\tilde{p}_4) = \begin{pmatrix} 2A_{\mathrm{fin}}(1^+,2^+,3^-,4^-) \\ 2A_{\mathrm{fin}}(1^+,3^-,4^-,2^+) \\ 2A_{\mathrm{fin}}(1^+,4^-,2^+,3^-) \\ B_{\mathrm{fin}}(1^+,2^+,3^-,4^-) \\ B_{\mathrm{fin}}(1^+,3^-,4^-,2^+) \\ B_{\mathrm{fin}}(1^+,4^-,2^+,3^-) \end{pmatrix},$$

$$\vec{C}_{+++-}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3,\tilde{p}_4) = \begin{pmatrix} 2A_{\mathrm{fin}}(1^+,2^+,3^+,4^-) \\ 2A_{\mathrm{fin}}(1^+,3^+,4^-,2^+) \\ 2A_{\mathrm{fin}}(1^+,4^-,2^+,3^+) \\ B_{\mathrm{fin}}(1^+,2^+,3^+,4^-) \\ B_{\mathrm{fin}}(1^+,3^+,4^-,2^+) \\ B_{\mathrm{fin}}(1^+,4^-,2^+,3^+) \end{pmatrix},$$

$$\vec{C}_{++++}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3,\tilde{p}_4) = \begin{pmatrix} 2A_{\mathrm{fin}}(1^+,2^+,3^+,4^+) \\ 2A_{\mathrm{fin}}(1^+,3^+,4^+,2^+) \\ 2A_{\mathrm{fin}}(1^+,4^+,2^+,3^+) \\ B_{\mathrm{fin}}(1^+,2^+,3^+,4^+) \\ B_{\mathrm{fin}}(1^+,3^+,4^+,2^+) \\ B_{\mathrm{fin}}(1^+,4^+,2^+,3^+) \end{pmatrix},$$

$$\vec{C}_{---+}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3,\tilde{p}_4) = \vec{C}_{+++-}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3,\tilde{p}_4)\Big|_{\langle..\rangle \leftrightarrow [..]},$$

$$\vec{C}_{----}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3,\tilde{p}_4) = \vec{C}_{++++}(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3,\tilde{p}_4)\Big|_{\langle..\rangle \leftrightarrow [..]}.$$

$$(5.203)$$

The last two coefficients follow from parity invariance. The factors of two in the first three entries of the coefficients come from combining the two color structures in the first three entries in Eq. (5.197) using charge conjugation invariance in Eq. (5.200).

The tree-level amplitudes are well known. At tree level, only the A amplitudes with two positive and two negative helicity gluons are nonzero. Because the A amplitudes correspond to a single-trace color structure, which possesses a cyclic symmetry, the corresponding partial amplitudes are invariant under the corresponding cyclic permutations of their arguments. Explicit expressions for the required amplitudes at tree level and one loop are given in Appendix B.4.1.

5.7.2 $pp \to$ 3 Jets

The four partonic channels $g\,q\bar{q}\,q'\bar{q}'$, $g\,q\bar{q}\,q\bar{q}$, $ggg\,q\bar{q}$, and $ggggg$ contribute to $pp \to$ 3 jets, which we discuss in turn. The one-loop partial amplitudes for the different partonic channels were calculated in Refs. [245, 246, 243]. Tree-level results for the helicity amplitudes for each partonic process are given in Appendix B.4.2.

$g\,q\bar{q}\,q'\bar{q}'$ and $g\,q\bar{q}\,q\bar{q}$

For the case of distinct quark flavors, $g\,q\bar{q}\,q'\bar{q}'$, the helicity basis consists of eight independent operators,

$$O^{a\,\bar{\alpha}\beta\bar{\gamma}\delta}_{\pm(+;+)} = \mathcal{B}^a_{1\pm}\, J^{\bar{\alpha}\beta}_{q\,23+}\, J^{\bar{\gamma}\delta}_{q'\,45+}\,,$$

$$O^{a\,\bar{\alpha}\beta\bar{\gamma}\delta}_{\pm(+;-)} = \mathcal{B}^a_{1\pm}\, J^{\bar{\alpha}\beta}_{q\,23+}\, J^{\bar{\gamma}\delta}_{q'\,45-}\,,$$

$$O^{a\,\bar{\alpha}\beta\bar{\gamma}\delta}_{\pm(-;+)} = \mathcal{B}^a_{1\pm}\, J^{\bar{\alpha}\beta}_{q\,23-}\, J^{\bar{\gamma}\delta}_{q'\,45+}\,,$$

$$O^{a\,\bar{\alpha}\beta\bar{\gamma}\delta}_{\pm(-;-)} = \mathcal{B}^a_{1\pm}\, J^{\bar{\alpha}\beta}_{q\,23-}\, J^{\bar{\gamma}\delta}_{q'\,45-}\,. \tag{5.204}$$

For identical quark flavors, $g\,q\bar{q}\,q\bar{q}$, the basis reduces to six independent helicity operators,

$$O^{a\,\bar{\alpha}\beta\bar{\gamma}\delta}_{\pm(++)} = \frac{1}{4}\mathcal{B}^a_{1\pm}\, J^{\bar{\alpha}\beta}_{23+}\, J^{\bar{\gamma}\delta}_{45+}\,,$$

289

$$O_{\pm(+-)}^{a\,\bar{\alpha}\beta\bar{\gamma}\delta} = \mathcal{B}_{1\pm}^{a}\,J_{23+}^{\bar{\alpha}\beta}\,J_{45-}^{\bar{\gamma}\delta}\,,$$

$$O_{\pm(--)}^{a\,\bar{\alpha}\beta\bar{\gamma}\delta} = \frac{1}{4}\,\mathcal{B}_{1\pm}^{a}\,J_{23-}^{\bar{\alpha}\beta}\,J_{45-}^{\bar{\gamma}\delta}\,. \tag{5.205}$$

In both cases we use the color basis

$$\bar{T}^{a\,\alpha\bar{\beta}\gamma\bar{\delta}} = 2T_{F}\Big(T_{\alpha\bar{\delta}}^{a}\,\delta_{\gamma\bar{\beta}},\ T_{\gamma\bar{\beta}}^{a}\,\delta_{\alpha\bar{\delta}},\ T_{\alpha\bar{\beta}}^{a}\,\delta_{\gamma\bar{\delta}},\ T_{\gamma\bar{\delta}}^{a}\,\delta_{\alpha\bar{\beta}}\Big). \tag{5.206}$$

The QCD helicity amplitudes for $g\,q\bar{q}\,q'\bar{q}'$ can be color decomposed into partial amplitudes in the color basis of Eq. (5.206) as

$$\mathcal{I}(g_1 q_2 \bar{q}_3 q'_4 \bar{q}'_5) = 2T_F \mathrm{i}\Big[T_{\alpha_2\bar{\alpha}_5}^{a_1}\delta_{\alpha_4\bar{\alpha}_3}A(1;2_q,3_{\bar{q}};4_{q'},5_{\bar{q}'}) + T_{\alpha_4\bar{\alpha}_3}^{a_1}\delta_{\alpha_2\bar{\alpha}_5}A(1;4_{q'},5_{\bar{q}'};2_q,3_{\bar{q}})$$

$$+ \frac{1}{N}T_{\alpha_2\bar{\alpha}_3}^{a_1}\delta_{\alpha_4\bar{\alpha}_5}B(1;2_q,3_{\bar{q}};4_{q'},5_{\bar{q}'}) + \frac{1}{N}T_{\alpha_4\bar{\alpha}_5}^{a_1}\delta_{\alpha_2\bar{\alpha}_3}B(1;4_{q'},5_{\bar{q}'};2_q,3_{\bar{q}})\Big]\,, \tag{5.207}$$

where we have used the symmetry $q\bar{q}\leftrightarrow q'\bar{q}'$, and inserted the factors of $1/N$ for later convenience. The amplitude vanishes when the quark and antiquark of the same flavor have the same helicity (both $+$ or both $-$), in accordance with the fact that the operators of Eq. (5.204) provide a complete basis of helicity operators. For identical quark flavors, the amplitudes can be written in terms of the amplitudes for the distinct flavor case as

$$\mathcal{I}(g_1 q_2 \bar{q}_3 q_4 \bar{q}_5) = \mathcal{I}(g_1 q_2 \bar{q}_3 q'_4 \bar{q}'_5) - \mathcal{I}(g_1 q_2 \bar{q}'_5 q'_4 \bar{q}_3)\,. \tag{5.208}$$

The Wilson coefficients for $g\,q\bar{q}\,q'\bar{q}'$ are then given by

$$\vec{C}_{+(+;+)}(\tilde{p}_1;\tilde{p}_2,\tilde{p}_3;\tilde{p}_4,\tilde{p}_5) = \begin{pmatrix} A_{\mathrm{fin}}(1^+;2_q^+,3_{\bar{q}}^-;4_{q'}^+,5_{\bar{q}'}^-) \\[4pt] A_{\mathrm{fin}}(1^+;4_{q'}^+,5_{\bar{q}'}^-;2_q^+,3_{\bar{q}}^-) \\[4pt] \frac{1}{N}B_{\mathrm{fin}}(1^+;2_q^+,3_{\bar{q}}^-;4_{q'}^+,5_{\bar{q}'}^-) \\[4pt] \frac{1}{N}B_{\mathrm{fin}}(1^+;4_{q'}^+,5_{\bar{q}'}^-;2_q^+,3_{\bar{q}}^-) \end{pmatrix}\,,$$

$$\vec{C}_{+(+;-)}(\tilde{p}_1;\tilde{p}_2,\tilde{p}_3;\tilde{p}_4,\tilde{p}_5) = \begin{pmatrix} A_{\text{fin}}(1^+;2_q^+,3_{\bar{q}}^-;4_{q'}^-,5_{\bar{q}'}^+) \\ A_{\text{fin}}(1^+;4_{q'}^-,5_{\bar{q}'}^+;2_q^+,3_{\bar{q}}^-) \\ \frac{1}{N}B_{\text{fin}}(1^+;2_q^+,3_{\bar{q}}^-;4_{q'}^-,5_{\bar{q}'}^+) \\ \frac{1}{N}B_{\text{fin}}(1^+;4_{q'}^-,5_{\bar{q}'}^+;2_q^+,3_{\bar{q}}^-) \end{pmatrix}, \qquad (5.209)$$

and for $g\,q\bar{q}\,q\bar{q}$ they are given in terms of the amplitudes A_{fin} and B_{fin} for $g\,q\bar{q}\,q'\bar{q}'$ by

$$\vec{C}_{+(++)}(\tilde{p}_1;\tilde{p}_2,\tilde{p}_3;\tilde{p}_4,\tilde{p}_5) = \begin{pmatrix} A_{\text{fin}}(1^+;2_q^+,3_{\bar{q}}^-;4_q^+,5_{\bar{q}}^-) - \frac{1}{N}B_{\text{fin}}(1^+;2_q^+,5_{\bar{q}}^-;4_q^+,3_{\bar{q}}^-) \\ A_{\text{fin}}(1^+;4_q^+,5_{\bar{q}}^-;2_q^+,3_{\bar{q}}^-) - \frac{1}{N}B_{\text{fin}}(1^+;4_q^+,3_{\bar{q}}^-;2_q^+,5_{\bar{q}}^-) \\ \frac{1}{N}B_{\text{fin}}(1^+;2_q^+,3_{\bar{q}}^-;4_q^+,5_{\bar{q}}^-) - A_{\text{fin}}(1^+;2_q^+,5_{\bar{q}}^-;4_q^+,3_{\bar{q}}^-) \\ \frac{1}{N}B_{\text{fin}}(1^+;4_q^+,5_{\bar{q}}^-;2_q^+,3_{\bar{q}}^-) - A_{\text{fin}}(1^+;4_q^+,3_{\bar{q}}^-;2_q^+,5_{\bar{q}}^-) \end{pmatrix},$$

$$\vec{C}_{+(+-)}(\tilde{p}_1;\tilde{p}_2,\tilde{p}_3;\tilde{p}_4,\tilde{p}_5) = \begin{pmatrix} A_{\text{fin}}(1^+;2_q^+,3_{\bar{q}}^-;4_q^-,5_{\bar{q}}^+) \\ A_{\text{fin}}(1^+;4_q^-,5_{\bar{q}}^+;2_q^+,3_{\bar{q}}^-) \\ \frac{1}{N}B_{\text{fin}}(1^+;2_q^+,3_{\bar{q}}^-;4_q^-,5_{\bar{q}}^+) \\ \frac{1}{N}B_{\text{fin}}(1^+;4_q^-,5_{\bar{q}}^+;2_q^+,3_{\bar{q}}^-) \end{pmatrix}. \qquad (5.210)$$

Charge conjugation invariance of QCD relates the Wilson coefficients,

$$\vec{C}_{\lambda(-;\pm)}(\tilde{p}_1;\tilde{p}_2,\tilde{p}_3;\tilde{p}_4,\tilde{p}_5) = \widehat{V}\vec{C}_{\lambda(+;\mp)}(\tilde{p}_1;\tilde{p}_3,\tilde{p}_2;\tilde{p}_5,\tilde{p}_4),$$

$$\vec{C}_{\lambda(--)}(\tilde{p}_1;\tilde{p}_2,\tilde{p}_3;\tilde{p}_4,\tilde{p}_5) = \widehat{V}\vec{C}_{\lambda(++)}(\tilde{p}_1;\tilde{p}_3,\tilde{p}_2;\tilde{p}_5,\tilde{p}_4), \qquad (5.211)$$

with

$$\widehat{V} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \qquad (5.212)$$

The remaining Wilson coefficients for a negative helicity gluon follow from parity invariance,

$$\vec{C}_{-(+;\pm)}(\tilde{p}_1;\tilde{p}_2,\tilde{p}_3;\tilde{p}_4,\tilde{p}_5)$$

291

$$= \vec{C}_{+(-;\mp)}(\tilde{p}_1; \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5)\Big|_{\langle..\rangle \leftrightarrow [..]} \, ,$$

$$\vec{C}_{-(++)}(\tilde{p}_1; \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5)$$

$$= \vec{C}_{+(--)}(\tilde{p}_1; \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5)\Big|_{\langle..\rangle \leftrightarrow [..]} \, . \tag{5.213}$$

Explicit expressions for all required partial amplitudes at tree level are given in Appendix B.4.2.

$ggg\,q\bar{q}$

For $ggg\,q\bar{q}$, we have a basis of eight independent helicity operators,

$$O^{abc\,\bar{\alpha}\beta}_{+++(\pm)} = \frac{1}{3!}\,\mathcal{B}^a_{1+}\,\mathcal{B}^b_{2+}\,\mathcal{B}^c_{3+}\,J^{\bar{\alpha}\beta}_{45\pm} \, ,$$

$$O^{abc\,\bar{\alpha}\beta}_{++-(\pm)} = \frac{1}{2}\,\mathcal{B}^a_{1+}\,\mathcal{B}^b_{2+}\,\mathcal{B}^c_{3-}\,J^{\bar{\alpha}\beta}_{45\pm} \, ,$$

$$O^{abc\,\bar{\alpha}\beta}_{--+(\pm)} = \frac{1}{2}\,\mathcal{B}^a_{1-}\,\mathcal{B}^b_{2-}\,\mathcal{B}^c_{3+}\,J^{\bar{\alpha}\beta}_{45\pm} \, ,$$

$$O^{abc\,\bar{\alpha}\beta}_{\quad(\pm)} = \frac{1}{3!}\mathcal{B}^a_{1-}\,\mathcal{B}^b_{2-}\,\mathcal{B}^c_{3-}\,J^{\bar{\alpha}\beta}_{45\pm} \, , \tag{5.214}$$

and we use the color basis

$$\bar{T}^{abc\,\alpha\bar{\beta}} = \begin{pmatrix} [T^aT^bT^c]_{\alpha\bar{\beta}} \\ [T^bT^cT^a]_{\alpha\bar{\beta}} \\ [T^cT^aT^b]_{\alpha\bar{\beta}} \\ [T^cT^bT^a]_{\alpha\bar{\beta}} \\ [T^aT^cT^b]_{\alpha\bar{\beta}} \\ [T^bT^aT^c]_{\alpha\bar{\beta}} \\ \mathrm{tr}[T^cT^a]T^b_{\alpha\bar{\beta}} \\ \mathrm{tr}[T^aT^b]T^c_{\alpha\bar{\beta}} \\ \mathrm{tr}[T^bT^c]T^a_{\alpha\bar{\beta}} \\ \mathrm{tr}[T^aT^bT^c]\delta_{\alpha\bar{\beta}} \\ \mathrm{tr}[T^cT^bT^a]\delta_{\alpha\bar{\beta}} \end{pmatrix}^T . \tag{5.215}$$

292

The color decomposition of the QCD helicity amplitudes into partial amplitudes using Eq. (5.215) is

$$\mathcal{I}\big(g_1 g_2 g_3\, q_4 \bar{q}_5\big) = i \sum_{\sigma \in S_3} \big[T^{a_{\sigma(1)}} T^{a_{\sigma(2)}} T^{a_{\sigma(3)}}\big]_{\alpha_4 \bar{\alpha}_5}$$

$$\times\, A(\sigma(1), \sigma(2), \sigma(3); 4_q, 5_{\bar{q}})$$

$$+\, i \sum_{\sigma \in S_3/Z_2} \mathrm{tr}\big[T^{a_{\sigma(1)}} T^{a_{\sigma(2)}}\big] T^{a_{\sigma(3)}}_{\alpha_4 \bar{\alpha}_5}$$

$$\times\, B(\sigma(1), \sigma(2), \sigma(3); 4_q, 5_{\bar{q}})$$

$$+\, i \sum_{\sigma \in S_3/Z_3} \mathrm{tr}\big[T^{a_{\sigma(1)}} T^{a_{\sigma(2)}} T^{a_{\sigma(3)}}\big] \delta_{\alpha_4 \bar{\alpha}_5}$$

$$\times\, C(\sigma(1), \sigma(2), \sigma(3); 4_q, 5_{\bar{q}}) , \tag{5.216}$$

from which we can read off the Wilson coefficients,

$$\vec{C}_{++\mp(+)}(\tilde{p}_1, \ldots; \tilde{p}_4, \tilde{p}_5) = \begin{pmatrix} A_{\mathrm{fin}}(1^+, 2^+, 3^\mp; 4_q^+, 5_{\bar{q}}^-) \\ A_{\mathrm{fin}}(2^+, 3^\mp, 1^+; 4_q^+, 5_{\bar{q}}^-) \\ A_{\mathrm{fin}}(3^\mp, 1^+, 2^+; 4_q^+, 5_{\bar{q}}^-) \\ A_{\mathrm{fin}}(3^\mp, 2^+, 1^+; 4_q^+, 5_{\bar{q}}^-) \\ A_{\mathrm{fin}}(1^+, 3^\mp, 2^+; 4_q^+, 5_{\bar{q}}^-) \\ A_{\mathrm{fin}}(2^+, 1^+, 3^\mp; 4_q^+, 5_{\bar{q}}^-) \\ B_{\mathrm{fin}}(3^\mp, 1^+, 2^+; 4_q^+, 5_{\bar{q}}^-) \\ B_{\mathrm{fin}}(1^+, 2^+, 3^\mp; 4_q^+, 5_{\bar{q}}^-) \\ B_{\mathrm{fin}}(2^+, 3^\mp, 1^+; 4_q^+, 5_{\bar{q}}^-) \\ C_{\mathrm{fin}}(1^+, 2^+, 3^\mp; 4_q^+, 5_{\bar{q}}^-) \\ C_{\mathrm{fin}}(3^\mp, 2^+, 1^+; 4_q^+, 5_{\bar{q}}^-) \end{pmatrix} . \tag{5.217}$$

Charge conjugation invariance of QCD relates the coefficients with opposite quark helicities,

$$\vec{C}_{\lambda_1 \lambda_2 \lambda_3 (-)}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5) = \widehat{V} \vec{C}_{\lambda_1 \lambda_2 \lambda_3 (+)}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3; \tilde{p}_5, \tilde{p}_4) ,$$

$$
\text{with} \quad \widehat{V} = \begin{pmatrix} 0_{3\times3} & 1_{3\times3} & & & \\ 1_{3\times3} & 0_{3\times3} & & & \\ & & 1_{3\times3} & & \\ & & & 0 & 1 \\ & & & 1 & 0 \end{pmatrix}, \tag{5.218}
$$

where $1_{n\times n}$ denotes the n-dimensional identity matrix and the empty entries are all zero. The remaining coefficients follow from parity invariance

$$
\begin{aligned}
&\vec{C}_{--+(\pm)}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5) \\
&\qquad = \vec{C}_{++-(\mp)}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5)\Big|_{\langle ..\rangle \leftrightarrow [..]}, \\
&\vec{C}_{---(\pm)}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5) \\
&\qquad = \vec{C}_{+++(\mp)}(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3; \tilde{p}_4, \tilde{p}_5)\Big|_{\langle ..\rangle \leftrightarrow [..]}.
\end{aligned} \tag{5.219}
$$

At tree level, the partial amplitudes are well known, and only the A amplitudes are nonzero. Furthermore, the partial amplitudes with all negative or all positive helicity gluons vanish. Combining the charge and parity relations of Eqs. (5.218) and (5.219), there are only three independent amplitudes at tree level, which we take to be $A(1^+, 2^+, 3^-; 4_q^+, 5_{\bar{q}}^-)$, $A(2^+, 3^-, 1^+; 4_q^+, 5_{\bar{q}}^-)$, and $A(3^-, 1^+, 2^+; 4_q^+, 5_{\bar{q}}^-)$. These amplitudes are given in Appendix B.4.2.

$ggggg$

For $ggggg$, the basis consists of six independent helicity operators,

$$
\begin{aligned}
O^{abcde}_{+++++} &= \frac{1}{5!} \, \mathcal{B}^a_{1+} \mathcal{B}^b_{2+} \mathcal{B}^c_{3+} \mathcal{B}^d_{4+} \mathcal{B}^e_{5+}, \\
O^{abcde}_{++++-} &= \frac{1}{4!} \, \mathcal{B}^a_{1+} \mathcal{B}^b_{2+} \mathcal{B}^c_{3+} \mathcal{B}^d_{4+} \mathcal{B}^e_{5-}, \\
O^{abcde}_{+++--} &= \frac{1}{2 \cdot 3!} \, \mathcal{B}^a_{1+} \mathcal{B}^b_{2+} \mathcal{B}^c_{3+} \mathcal{B}^d_{4-} \mathcal{B}^e_{5-}, \\
O^{abcde}_{---++} &= \frac{1}{2 \cdot 3!} \, \mathcal{B}^a_{1-} \mathcal{B}^b_{2-} \mathcal{B}^c_{3-} \mathcal{B}^d_{4+} \mathcal{B}^e_{5+},
\end{aligned}
$$

$$O^{abcde}_{----+} = \frac{1}{4!}\, \mathcal{B}^a_{1-}\mathcal{B}^b_{2-}\mathcal{B}^c_{3-}\mathcal{B}^d_{4-}\mathcal{B}^e_{5+}\,,$$

$$O^{abcde}_{-----} = \frac{1}{5!}\, \mathcal{B}^a_{1-}\mathcal{B}^b_{2-}\mathcal{B}^c_{3-}\mathcal{B}^d_{4-}\mathcal{B}^e_{5-}\,. \tag{5.220}$$

As before, we only need one operator for each number of positive and negative helicities. We use the color basis

$$\bar{T}^{abcde} = \frac{1}{2 \cdot 2T_F}\begin{pmatrix} \mathrm{tr}[abcde] - \mathrm{tr}[edcba] \\ \mathrm{tr}[acdeb] - \mathrm{tr}[bedca] \\ \mathrm{tr}[acbed] - \mathrm{tr}[debca] \\ \mathrm{tr}[abced] - \mathrm{tr}[decba] \\ \mathrm{tr}[abdec] - \mathrm{tr}[cedba] \\ \mathrm{tr}[acbde] - \mathrm{tr}[edbca] \\ \mathrm{tr}[adceb] - \mathrm{tr}[becda] \\ \mathrm{tr}[adcbe] - \mathrm{tr}[ebcda] \\ \mathrm{tr}[aebdc] - \mathrm{tr}[cdbea] \\ \mathrm{tr}[abdce] - \mathrm{tr}[ecdba] \\ \mathrm{tr}[aecbd] - \mathrm{tr}[dbcea] \\ \mathrm{tr}[acebd] - \mathrm{tr}[dbeca] \\ (\mathrm{tr}[ced] - \mathrm{tr}[dec])\mathrm{tr}[ab] \\ (\mathrm{tr}[abe] - \mathrm{tr}[eba])\mathrm{tr}[cd] \\ (\mathrm{tr}[acd] - \mathrm{tr}[dca])\mathrm{tr}[be] \\ (\mathrm{tr}[bec] - \mathrm{tr}[ceb])\mathrm{tr}[ad] \\ (\mathrm{tr}[adb] - \mathrm{tr}[bda])\mathrm{tr}[ce] \\ (\mathrm{tr}[ace] - \mathrm{tr}[eca])\mathrm{tr}[bd] \\ (\mathrm{tr}[bdc] - \mathrm{tr}[cdb])\mathrm{tr}[ae] \\ (\mathrm{tr}[aed] - \mathrm{tr}[dea])\mathrm{tr}[bc] \\ (\mathrm{tr}[acb] - \mathrm{tr}[bca])\mathrm{tr}[de] \\ (\mathrm{tr}[bed] - \mathrm{tr}[deb])\mathrm{tr}[ac] \end{pmatrix}^T, \tag{5.221}$$

where we have used the shorthand notation

$$\text{tr}[ab\cdots cd] = \text{tr}[T^a T^b \cdots T^c T^d].$$

(5.222)

A priori, there are twice as many color structures as in Eq. (5.221) with a relative plus sign instead of a minus sign between the two traces. Under charge conjugation, the operators transform as

$$C\, O^{abcde}_{\lambda_1 \lambda_2 \lambda_3 \lambda_4 \lambda_5}\, \bar{T}^{abcde}\, C = -O^{edcba}_{\lambda_1 \lambda_2 \lambda_3 \lambda_4 \lambda_5}(\bar{T}^{abcde}).$$

(5.223)

Therefore, charge conjugation invariance implies for the Wilson coefficients

$$C^{abcde}_{\lambda_1 \lambda_2 \lambda_3 \lambda_4 \lambda_5} = -C^{edcba}_{\lambda_1 \lambda_2 \lambda_3 \lambda_4 \lambda_5},$$

(5.224)

and hence these additional color structures cannot appear at any order in perturbation theory, either through matching or renormalization group evolution.

The color decomposition of the QCD amplitude into partial amplitudes using the color basis of Eq. (5.221) is

$$
\begin{aligned}
\mathcal{I}&(g_1 g_2 g_3 g_4 g_5)\\
&= \frac{i}{2T_F}\Bigg[\sum_{\sigma \in S_5/Z_5} \text{tr}[a_{\sigma(1)} a_{\sigma(2)} a_{\sigma(3)} a_{\sigma(4)} a_{\sigma(5)}]\\
&\quad \times A\big(\sigma(1), \sigma(2), \sigma(3), \sigma(4), \sigma(5)\big)\\
&\quad + \sum_{\sigma \in S_5/(Z_3 \times Z_2)} \text{tr}[a_{\sigma(1)} a_{\sigma(2)} a_{\sigma(3)}]\, \text{tr}[a_{\sigma(4)} a_{\sigma(5)}]\\
&\quad \times B\big(\sigma(1), \sigma(2), \sigma(3), \sigma(4), \sigma(5)\big)\Bigg],
\end{aligned}
$$

(5.225)

296

from which we obtain the Wilson coefficients

$$
\vec{C}_{+++--}(\tilde{p}_1,\dots,\tilde{p}_5) = 2
\begin{pmatrix}
A_{\text{fin}}(1^+,2^+,3^+,4^-,5^-) \\
A_{\text{fin}}(1^+,3^+,4^-,5^-,2^+) \\
A_{\text{fin}}(1^+,3^+,2^+,5^-,4^-) \\
A_{\text{fin}}(1^+,2^+,3^+,5^-,4^-) \\
A_{\text{fin}}(1^+,2^+,4^-,5^-,3^+) \\
A_{\text{fin}}(1^+,3^+,2^+,4^-,5^-) \\
A_{\text{fin}}(1^+,4^-,3^+,5^-,2^+) \\
A_{\text{fin}}(1^+,4^-,3^+,2^+,5^-) \\
A_{\text{fin}}(1^+,5^-,2^+,4^-,3^+) \\
A_{\text{fin}}(1^+,2^+,4^-,3^+,5^-) \\
A_{\text{fin}}(1^+,5^-,3^+,2^+,4^-) \\
A_{\text{fin}}(1^+,3^+,5^-,2^+,4^-) \\
B_{\text{fin}}(3^+,5^-,4^-,1^+,2^+) \\
B_{\text{fin}}(1^+,2^+,5^-,3^+,4^-) \\
B_{\text{fin}}(1^+,3^+,4^-,2^+,5^-) \\
B_{\text{fin}}(2^+,5^-,3^+,1^+,4^-) \\
B_{\text{fin}}(1^+,4^+,2^-,3^+,5^-) \\
B_{\text{fin}}(1^+,3^+,5^-,2^+,4^-) \\
B_{\text{fin}}(2^+,4^-,3^+,1^+,5^-) \\
B_{\text{fin}}(1^+,5^-,4^-,2^+,3^+) \\
B_{\text{fin}}(1^+,3^+,2^+,4^-,5^-) \\
B_{\text{fin}}(2^+,5^-,4^-,1^+,3^+)
\end{pmatrix},
$$

$$
\vec{C}_{----\pm}(\tilde{p}_1,\dots,\tilde{p}_5) = \vec{C}_{++++\mp}(\tilde{p}_1,\dots,\tilde{p}_5)\Big|_{\langle..\rangle\leftrightarrow[..]} ,
$$

$$
\vec{C}_{---++}(\tilde{p}_1,\dots,\tilde{p}_5) = \vec{C}_{+++--}(\tilde{p}_1,\dots,\tilde{p}_5)\Big|_{\langle..\rangle\leftrightarrow[..]} . \tag{5.226}
$$

For brevity, we have not written out the coefficients $\vec{C}_{++++-}$ and $\vec{C}_{+++++}$. They have

exactly the same structure as $\vec{C}_{+++--}$ with the replacements $4^- \to 4^+$ and $4^-, 5^- \to 4^+, 5^+$, respectively, in the arguments of the helicity amplitudes. The remaining Wilson coefficients are given by parity invariance as shown. The overall factor of two comes from combining the two color structures in Eq. (5.221), which are related by charge conjugation.

At tree level, all the B amplitudes vanish, as do all the amplitudes in $\vec{C}_{++++\pm}$ and $\vec{C}_{----\mp}$. By the parity relations given in Eq. (5.226), only the A amplitudes in $\vec{C}_{+++--}$ are then required for the tree-level matching. Since these amplitudes correspond to single trace color structures, which posses a cyclic symmetry, the required partial amplitudes are invariant under the corresponding cyclic permutations of their arguments. Therefore, at tree level, there are only two independent amplitudes, which we take to be $A_{\mathrm{fin}}(1^+, 2^+, 3^+, 4^-, 5^-)$ and $A_{\mathrm{fin}}(1^+, 2^+, 4^-, 3^+, 5^-)$. These are given in Appendix B.4.2. Simplifications also occur at one loop, since the B amplitudes can be expressed in terms of sums of permutations of the A amplitudes [229, 230].

5.8 Renormalization Group Evolution

In this section, we discuss the renormalization group evolution (RGE) of the Wilson coefficients. We start with a general discussion and give the solution of the RGE to all orders in perturbation theory. For completeness, we also explicitly derive the (known) anomalous dimension at one loop. To discuss the RGE, it is convenient to consider the operators $\vec{O}^\dagger$ in Eq. (5.54), which are vectors in color space. Lastly, we give explicit results, in a manifestly crossing symmetric form, for the relevant color mixing matrices for the color bases we have used in the previous sections. Since the operators' renormalization is independent of their helicity structure, we drop all helicity labels throughout this section for notational simplicity.

5.8.1 General Discussion

The renormalization of the hard scattering in SCET can either be carried out as operator renormalization, where the relation between bare and renormalized matrix elements is

$\langle \vec{O}^\dagger \rangle^{\text{bare}} = Z_\xi^{-n_q/2} Z_A^{-n_g/2} \langle \vec{O}^\dagger \rangle^{\text{ren}} \widehat{Z}_O$, or with coefficient renormalization where $\langle \vec{O}^\dagger \rangle^{\text{bare}} \vec{C}^{\text{bare}} = Z_\xi^{n_q/2} Z_A^{n_g/2} \langle \vec{O}^\dagger \rangle^{\text{bare}} \widehat{Z}_C \vec{C}^{\text{ren}}$. The relationship between the two is $\widehat{Z}_C = \widehat{Z}_O^{-1}$. Here Z_ξ and Z_A are the wave-function renormalizations of the SCET collinear quark and gluon fields ξ_n and A_n, defined in Sec. 5.2.2, and

$$n_g = n_g^+ + n_g^- \,, \qquad n_q = n_q^+ + n_q^- \tag{5.227}$$

are the total number of quark and gluon helicity fields in the operator (recall that there are two quark fields in each of the fermionic helicity currents). The UV divergences for $\langle \vec{O}^\dagger \rangle^{\text{bare}}$ are given in terms of a local product (as opposed to a convolution over label momenta), since we are working at leading power where the operators contain a single field per collinear sector.

Let us consider more explicitly how the renormalization works at one loop. The counterterm Feynman rule at this order is

$$\langle \vec{O}^\dagger \rangle^{\text{tree}} \left(Z_\xi^{n_q/2} Z_A^{n_g/2} \widehat{Z}_C - 1 \right). \tag{5.228}$$

At one loop, the UV divergences of $\langle \vec{O}^\dagger \rangle^{\text{bare}}$ are proportional to the tree-level matrix element as $\langle \vec{O}^\dagger \rangle^{\text{tree}} \widehat{D}$, where $\widehat{D}$ is a matrix in color space, which denotes the $1/\epsilon^2$ and $1/\epsilon$ UV divergences (with μ defined in the $\overline{\text{MS}}$ scheme) of the bare matrix element. The counterterm has to cancel these UV divergences so

$$\langle \vec{O}^\dagger \rangle^{\text{tree}} \left(Z_\xi^{n_q/2} Z_A^{n_g/2} \widehat{Z}_C - 1 \right) = - \langle \vec{O}^\dagger \rangle^{\text{tree}} \widehat{D} \,, \tag{5.229}$$

which fixes $\widehat{Z}_C$ at one loop.

Next consider the renormalization group equations, working to all orders in α_s. As usual, the μ independence of the bare operator implies the renormalization group equation for the Wilson coefficient

$$\mu \frac{\mathrm{d}\vec{C}(\mu)}{\mathrm{d}\mu} = \widehat{\gamma}_C(\mu) \, \vec{C}(\mu) \,, \tag{5.230}$$

where the anomalous dimension matrix is defined as

$$\widehat{\gamma}_C(\mu) = -\widehat{Z}_C^{-1}(\mu)\left[\frac{\mathrm{d}}{\mathrm{d}\ln\mu}\,\widehat{Z}_C(\mu)\right].$$

(5.231)

The solution of the RGE in Eq. (5.230) can be written as

$$\vec{C}(\mu) = \widehat{U}(\mu_0,\mu)\,\vec{C}(\mu_0),$$

(5.232)

with the evolution matrix

$$\widehat{U}(\mu_0,\mu) = \mathcal{P}\exp\left[\int_{\ln\mu_0}^{\ln\mu}\mathrm{d}\ln\mu'\,\widehat{\gamma}_C(\mu')\right].$$

(5.233)

Here, $\mathcal{P}$ denotes path ordering along increasing μ, and $\mu > \mu_0$. The path ordering is necessary since $\widehat{\gamma}_C(\mu)$ is a matrix in color space.

The anomalous dimension matrix has the general form

$$\widehat{\gamma}_C(\mu) = \Gamma_{\mathrm{cusp}}[\alpha_s(\mu)]\,\widehat{\Delta}(\mu^2) + \widehat{\gamma}[\alpha_s(\mu)],$$

(5.234)

where Γ_{cusp} is the cusp anomalous dimension and $\widehat{\Delta}(\mu^2)$ is a process-dependent mixing matrix in color space, which does not depend on α_s. Its μ dependence is given by

$$\widehat{\Delta}(\mu^2) = \mathbf{1}(n_g C_A + n_q C_F)\ln\left(\frac{\mu_0}{\mu}\right) + \widehat{\Delta}(\mu_0^2),$$

(5.235)

which will be demonstrated explicitly at one loop in Sec. 5.8.2. We can then perform the integral in Eq. (5.233) by using the running of the coupling, $\mathrm{d}\alpha_s(\mu)/\mathrm{d}\ln\mu = \beta(\alpha_s)$, to switch variables from $\ln\mu$ to α_s. We find

$$\widehat{U}(\mu_0,\mu) = e^{-(n_g C_A + n_q C_F)K_\Gamma(\mu_0,\mu)}$$

$$\times\,\bar{\mathcal{P}}_{\alpha_s}\exp\left[\eta_\Gamma(\mu_0,\mu)\,\widehat{\Delta}(\mu_0^2) + \widehat{K}_\gamma(\mu_0,\mu)\right],$$

(5.236)

300

where $\bar{\mathcal{P}}_{\alpha_s}$ now denotes path ordering along decreasing α_s, with $\alpha_s(\mu) < \alpha_s(\mu_0)$, and

$$
K_\Gamma(\mu_0, \mu) = \int_{\alpha_s(\mu_0)}^{\alpha_s(\mu)} d\alpha_s \, \frac{\Gamma_{\text{cusp}}(\alpha_s)}{\beta(\alpha_s)} \int_{\alpha_s(\mu_0)}^{\alpha_s} d\alpha_s' \, \frac{1}{\beta(\alpha_s')} \,,
$$
$$
\eta_\Gamma(\mu_0, \mu) = \int_{\alpha_s(\mu_0)}^{\alpha_s(\mu)} d\alpha_s \, \frac{\Gamma_{\text{cusp}}(\alpha_s)}{\beta(\alpha_s)} \,,
$$
$$
\widehat{K}_\gamma(\mu_0, \mu) = \int_{\alpha_s(\mu_0)}^{\alpha_s(\mu)} d\alpha_s \, \frac{\widehat{\gamma}(\alpha_s)}{\beta(\alpha_s)} \,.
\tag{5.237}
$$

Up to two loops, the noncusp piece $\widehat{\gamma}(\alpha_s)$ in Eq. (5.234) is proportional to the identity operator [376, 377]

$$
\widehat{\gamma}(\alpha_s) = (n_q \gamma_C^q + n_g \gamma_C^g) \mathbf{1} \,.
\tag{5.238}
$$

In this case, the evolution factor simplifies to

$$
\widehat{U}(\mu_0, \mu) = e^{-(n_g C_A + n_q C_F) K_\Gamma(\mu_0, \mu) + K_\gamma(\mu_0, \mu)}
$$
$$
\times \exp\left[\eta_\Gamma(\mu_0, \mu) \, \widehat{\Delta}(\mu_0^2)\right] \,.
\tag{5.239}
$$

Starting at three loops the noncusp anomalous dimension is not color diagonal, and starts to depend on a conformal cross ratio built from factors of $p_i \cdot p_j$ [378]. (For earlier work beyond two loops see Refs. [379, 380, 381, 382, 383, 384, 385, 386, 387]. The result of Ref. [378] implies that the conjectured all-order dipole color structure in Refs. [382, 383] is violated.)

The evolution factors $K_\Gamma(\mu_0, \mu)$, and $\eta_\Gamma(\mu_0, \mu)$ are universal. Explicit expressions for the integrals in Eq. (5.237) to NNLL order, together with the required coefficients for Γ_{cusp} and the β function to three loops, are given for reference in Appendix B.5.

5.8.2 One-loop Anomalous Dimension

The anomalous dimension $\widehat{\gamma}_C(\mu)$ is process dependent. In this subsection, we derive its general form at one loop. The anomalous dimension of the operators is determined from the UV divergences in the effective theory. The relevant one-loop diagrams in SCET are shown

301

(a)

(b)
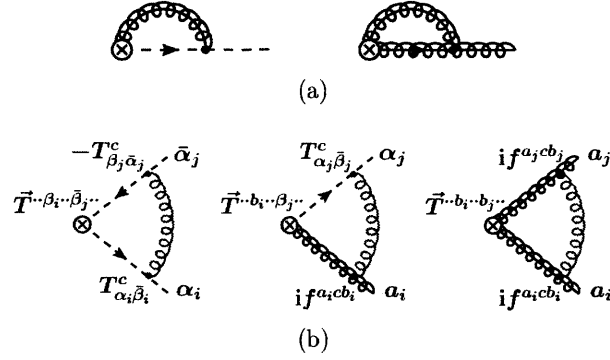
Figure 5-1: (a) Collinear one-loop diagrams. (b) Soft one-loop diagrams connecting two fields i and j in the operator.

in Fig. 6-4. In pure dimensional regularization the UV and IR divergences cancel such that the bare results for the loop diagrams vanish. To extract the UV divergences, we regulate the IR divergences by taking the external particles off shell with $p_i^2 = p_{i\perp}^2 \neq 0$.

Since all fields in the operators correspond to distinct collinear directions, the collinear loop diagrams in Fig. 5-1a only involve one external line at a time. Different external lines can only interact through the exchange of a soft gluon, shown by the diagrams in Fig. 5-1b.

When expressing our results, we use the notation [see Eq. (5.73)]

$$L_{i\perp} = \ln\left(-\frac{p_{i\perp}^2}{\mu^2}\right), \qquad L_{ij} = \ln\left(-\frac{s_{ij}}{\mu^2} - i0\right), \tag{5.240}$$

where $s_{ij} = 2p_i \cdot p_j$.

First, we recall the wave function renormalization constants. In Feynman gauge at one loop,

$$Z_\xi = 1 - \frac{\alpha_s}{4\pi}\frac{1}{\epsilon}(C_F + \cdots),$$
$$Z_A = 1 + \frac{\alpha_s}{4\pi}\frac{1}{\epsilon}(\beta_0 - 2C_A + \cdots), \tag{5.241}$$

where $\beta_0 = 11/3\,C_A - 4/3\,T_F n_f$ is the one-loop beta function coefficient [see Eq. (B.120)], and n_f is the number of considered quark flavors. Here and below, the ellipses denote possible UV-finite terms, which are irrelevant for our discussion here. (Using the on-shell scheme for

302

wave function renormalization, the Z_i contain UV-finite pieces, see Appendix B.7.)

The collinear diagrams in Fig. 5-1a contribute

$$I_c^q = I_c^{\bar{q}} = \frac{\alpha_s C_F}{4\pi} \left(\frac{2}{\epsilon^2} + \frac{2}{\epsilon} - \frac{2}{\epsilon} L_{i\perp} + \cdots \right) \langle \vec{O}^\dagger \rangle^{\text{tree}},$$
$$I_c^g = \frac{\alpha_s C_A}{4\pi} \left(\frac{2}{\epsilon^2} + \frac{1}{\epsilon} - \frac{2}{\epsilon} L_{i\perp} + \cdots \right) \langle \vec{O}^\dagger \rangle^{\text{tree}}, \qquad (5.242)$$

where I_c^i denotes the result of the diagram for an external leg of type i, either quark or gluon.

The soft diagrams in Fig. 5-1b differ from each other only in their color structure. The result of the diagram connecting particles i and j (with $i \neq j$) is given by

$$I_s^{ij} = \frac{\alpha_s}{4\pi} \left(\frac{2}{\epsilon^2} + \frac{2}{\epsilon} L_{ij} - \frac{2}{\epsilon} L_{i\perp} - \frac{2}{\epsilon} L_{j\perp} + \cdots \right) \langle \vec{O}^\dagger \rangle^{\text{tree}} \hat{t}_i^c \hat{t}_j^c, \qquad (5.243)$$

where $\hat{t}_i^c$ and $\hat{t}_j^c$ are matrices in color space. From Eqs. (5.242) and (5.243) we see explicitly that the operators only mix with respect to the color structure, with no mixing between operators with distinct helicities.

The action of the matrix $\hat{t}_i^c$ on the color space is to insert a generator acting on the color index of the ith particle, i.e.,

$$(\bar{T}\,\hat{t}_i^c)^{\cdots\alpha_i\cdots} = T_{\alpha_i\bar{\beta}_i}^c \, \bar{T}^{\cdots\beta_i\cdots},$$
$$(\bar{T}\,\hat{t}_i^c)^{\cdots\bar{\alpha}_i\cdots} = -\bar{T}^{\cdots\bar{\beta}_i\cdots} \, T_{\beta_i\bar{\alpha}_i}^c,$$
$$(\bar{T}\,\hat{t}_i^c)^{\cdots a_i\cdots} = i f^{a_i c b_i} \, \bar{T}^{\cdots b_i\cdots}, \qquad (5.244)$$

for quarks, antiquarks, and gluons, respectively. Our $\hat{t}_i^c$ is identical to what is usually denoted as $\mathbf{T}_i$ in the notation of Refs. [388, 389].

To give an explicit example, consider $gg\,q\bar{q}$. Then, for quark $i = 3$ and antiquark $j = 4$ we have

$$\vec{O}^\dagger \, \hat{t}_3^c \, \hat{t}_4^c = O^{a_1 a_2 \bar{\alpha}_3 \alpha_4} \, (\bar{T}\,\hat{t}_3^c\,\hat{t}_4^c)^{a_1 a_2 \alpha_3 \bar{\alpha}_4}$$
$$= O^{a_1 a_2 \bar{\alpha}_3 \alpha_4} \, T_{\alpha_3 \bar{\beta}_3}^c \, (-T_{\beta_4 \bar{\alpha}_4}^c) \bar{T}^{a_1 a_2 \beta_3 \bar{\beta}_4}, \qquad (5.245)$$

303

while for gluon $i = 1$ and quark $j = 3$,

$$\vec{O}^{\dagger}\,\hat{t}_1^c\,\hat{t}_3^c = O^{a_1 a_2 \bar{\alpha}_3 \alpha_4}\,\mathrm{i} f^{a_1 c b_1} T^c_{\alpha_3 \bar{\beta}_3}\,\bar{T}^{b_1 a_2 \beta_3 \bar{\alpha}_4}\,. \tag{5.246}$$

Plugging in the explicit basis in Eq. (5.192) and using the relations in Appendix B.1.2, we can rewrite the resulting color structures above in terms of the basis in Eq. (5.192), which yields

$$\hat{t}_3^c\,\hat{t}_4^c = -\begin{pmatrix} C_F - \frac{1}{2}C_A & 0 & 0 \\ 0 & C_F - \frac{1}{2}C_A & 0 \\ T_F & T_F & C_F \end{pmatrix},$$

$$\hat{t}_1^c\,\hat{t}_3^c = -\begin{pmatrix} \frac{1}{2}C_A & 0 & T_F \\ 0 & 0 & -T_F \\ 0 & -T_F & 0 \end{pmatrix}. \tag{5.247}$$

The other combinations are computed analogously.

In general, one can easily see that for $i = j$

$$\bar{T}^{a_1 \cdots a_n}\,\hat{t}_i^c\,\hat{t}_i^c = C_i\,\bar{T}^{a_1 \cdots a_n}\,, \tag{5.248}$$

where $C_i = C_F$ for quarks and $C_i = C_A$ for gluons. By construction, the color basis $\bar{T}^{a_1 \cdots a_n}$ conserves color, because each index corresponds to an external particle. Since $\hat{t}_i^c$ measures the color charge of the ith particle, color conservation implies

$$\bar{T}^{a_1 \cdots a_n}\left(\sum_{i=1}^{n} \hat{t}_i^c\right) = 0\,. \tag{5.249}$$

As a simple example, consider $gq\bar{q}$ for which $\bar{T}^{a_1 \alpha_2 \bar{\alpha}_3} \equiv T^{a_1}_{\alpha_2 \bar{\alpha}_3}$. In this case, Eq. (5.249) gives

$$\mathrm{i} f^{a_1 c b_1} T^{b_1}_{\alpha_2 \bar{\alpha}_3} + T^c_{\alpha_2 \bar{\beta}_2} T^{a_1}_{\beta_2 \bar{\alpha}_3} - T^{a_1}_{\alpha_2 \bar{\beta}_3} T^c_{\beta_3 \bar{\alpha}_3}$$

$$= \left(\mathrm{i} f^{a_1 c b_1} T^{b_1} + [T^c, T^{a_1}]\right)_{\alpha_2 \bar{\alpha}_3} = 0\,. \tag{5.250}$$

The total bare one-loop matrix element is given by summing Eq. (5.242) for each external particle and Eq. (5.243) for each pair of distinct particles. The infrared logarithms $L_{i\perp}$ have to drop out in the sum of all UV-divergent contributions. To see that this is indeed the case, we can use Eq. (5.248) to rewrite the collinear contributions. Then, the sum of all $L_{i\perp}$ terms is proportional to

$$
\begin{aligned}
\langle\vec{O}^{\dagger}\rangle^{\text{tree}} &\left[\sum_i L_{i\perp}\hat{t}_i^c\hat{t}_i^c + \sum_{i<j}(L_{i\perp}+L_{j\perp})\,\hat{t}_i^c\hat{t}_j^c\right] \\
&= \langle\vec{O}^{\dagger}\rangle^{\text{tree}}\left(\sum_i L_{i\perp}\hat{t}_i^c\hat{t}_i^c + \sum_{i\neq j}L_{i\perp}\hat{t}_i^c\hat{t}_j^c\right) \\
&= \langle\vec{O}^{\dagger}\rangle^{\text{tree}}\left(\sum_i L_{i\perp}\hat{t}_i^c\right)\left(\sum_j \hat{t}_j^c\right) = 0\,,
\end{aligned}
\tag{5.251}
$$

where in the last step we used Eq. (5.249). For the same reason the $1/\epsilon^2$ poles in the soft diagrams cancel against half of the $1/\epsilon^2$ poles in the collinear diagrams. The remaining UV-divergent part of the matrix element is given by

$$
\begin{aligned}
\langle\vec{O}^{\dagger}\rangle^{\text{tree}}\widehat{D} = \langle\vec{O}^{\dagger}\rangle^{\text{tree}}\frac{\alpha_s}{4\pi}\bigg[& n_g C_A\left(\frac{1}{\epsilon^2}+\frac{1}{\epsilon}\right) \\
& + n_q C_F\left(\frac{1}{\epsilon^2}+\frac{2}{\epsilon}\right) - \frac{2}{\epsilon}\widehat{\Delta}(\mu^2)\bigg]\,,
\end{aligned}
\tag{5.252}
$$

where the color mixing matrix is given by

$$
\widehat{\Delta}(\mu^2) = -\sum_{i<j}\hat{t}_i^c\hat{t}_j^c\,L_{ij}\,.
\tag{5.253}
$$

Combining this result with the identities in Eqs. (5.248) and (5.249), we can easily check that the μ dependence of $\widehat{\Delta}(\mu^2)$ is as in Eq. (5.235):

$$
\begin{aligned}
\widehat{\Delta}(\mu^2) - \widehat{\Delta}(\mu_0^2) &= -2\sum_{i<j}\hat{t}_i^c\hat{t}_j^c\,\ln\left(\frac{\mu_0}{\mu}\right) = \sum_i \hat{t}_i^c\hat{t}_i^c\,\ln\left(\frac{\mu_0}{\mu}\right) \\
&= \mathbf{1}(n_g C_A + n_q C_F)\ln\left(\frac{\mu_0}{\mu}\right)\,.
\end{aligned}
\tag{5.254}
$$

305

We can now compute the anomalous dimension of the operators. From Eqs. (5.229) and (5.252), we find at one loop

$$\widehat{Z}_C = \mathbf{1} - \widehat{D} - \mathbf{1}\left[\frac{n_g}{2}\left(Z_A - 1\right) + \frac{n_q}{2}\left(Z_\xi - 1\right)\right],$$ (5.255)

which using Eq. (5.231) yields the one-loop anomalous dimension

$$\widehat{\gamma}_C(\mu) = \frac{\alpha_s(\mu)}{4\pi}\left[4\widehat{\Delta}(\mu^2) - \mathbf{1}(n_g\beta_0 + n_q\,3C_F)\right].$$ (5.256)

The coefficient of 4 in front of $\widehat{\Delta}(\mu^2)$ is the one-loop cusp anomalous dimension coefficient [see Eq. (B.121)]. The remaining terms determine the noncusp $\widehat{\gamma}(\alpha_s)$ in Eq. (5.234) at one loop,

$$\widehat{\gamma}(\alpha_s) = -\frac{\alpha_s}{4\pi}\left(n_g\beta_0 + n_q\,3C_F\right)\mathbf{1}.$$ (5.257)

5.8.3 Mixing Matrices

In this section, we give explicit expressions for the mixing matrices for the color bases used in Secs. 5.5, 5.6, and 5.7. For simplicity, we only give explicit expressions for up to four partons, but allow for additional colorless particles, such as a Higgs or vector boson. The matrices are straightforward to evaluate using the color relations in Appendix B.1.2, but become rather lengthy for more than four partons, due to the large number of allowed color structures, and are more easily evaluated in an automated way (see for example Ref. [390]). For convenience, we introduce the following shorthand notation for sums and differences of logarithms L_{ij},

$$L_{ij\cdot kl\cdots} = L_{ij} + L_{kl} + \dots,$$

$$L_{ij\cdots/(kl\cdots)} = (L_{ij\cdots}) - (L_{kl\cdots}),$$ (5.258)

with $L_{ij} = \ln(-s_{ij}/\mu^2 - i0)$ as defined in Eq. (5.73).

Pure Gluon Mixing Matrices

For gg and ggg in the bases used in Eq. (5.105) and Eqs. (5.110) and (5.157), we have

$$\widehat{\Delta}_{gg}(\mu^2) = C_A\, L_{12}\,, \qquad \widehat{\Delta}_{ggg}(\mu^2) = \frac{1}{2} C_A\, L_{12\cdot13\cdot23} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{5.259}$$

For $gggg$ in the basis used in Eqs. (5.130) and (5.197), we have

$$
\widehat{\Delta}_{gggg}(\mu^2)
$$

$$
= \begin{pmatrix}
\frac{1}{2}C_A L_{12\cdot14\cdot23\cdot34} & 0 & 0 & 2T_F L_{14\cdot23/(13\cdot24)} & 0 & 2T_F L_{12\cdot34/(13\cdot24)} \\
0 & \frac{1}{2}C_A L_{12\cdot13\cdot24\cdot34} & 0 & 2T_F L_{13\cdot24/(14\cdot23)} & 2T_F L_{12\cdot34/(14\cdot23)} & 0 \\
0 & 0 & \frac{1}{2}C_A L_{13\cdot14\cdot23\cdot24} & 0 & 2T_F L_{14\cdot23/(12\cdot34)} & 2T_F L_{13\cdot24/(12\cdot34)} \\
T_F L_{12\cdot34/(13\cdot24)} & T_F L_{12\cdot34/(14\cdot23)} & 0 & C_A L_{12\cdot34} & 0 & 0 \\
0 & T_F L_{13\cdot24/(14\cdot23)} & T_F L_{13\cdot24/(12\cdot34)} & 0 & C_A L_{13\cdot24} & 0 \\
T_F L_{14\cdot23/(13\cdot24)} & 0 & T_F L_{14\cdot23/(12\cdot34)} & 0 & 0 & C_A L_{14\cdot23}
\end{pmatrix}.
$$

$$\tag{5.260}$$

For our color bases formed from multitrace color structures, the structure of the mixing matrices is simple. Since the mixing matrices are determined by single gluon exchange, cyclicity is maintained, and all that can occur in the mixing is that a single trace splits into two or two traces recombine into one. For example, the color structure $\mathrm{tr}[T^a T^b T^c T^d]$ can only mix with

$$\mathrm{tr}[T^a T^b T^c T^d]\,, \qquad \mathrm{tr}[T^a T^b]\,\mathrm{tr}[T^c T^d]\,, \qquad \text{and} \qquad \mathrm{tr}[T^d T^a]\,\mathrm{tr}[T^b T^c]\,. \tag{5.261}$$

Therefore, although the mixing matrices quickly get large as the number of color structures grows, their structure remains relatively simple. (An alternative approach to the organization of the anomalous dimensions for a large number of partons has been given in Ref. [391].) For the dijet case, i.e., in the absence of additional colorless particles, the kinematics simplifies

307

to

$$s = s_{12} = s_{34}\,, \qquad t = s_{13} = s_{24}\,, \qquad u = s_{14} = s_{23}\,, \tag{5.262}$$

and these matrices were given in Ref. [335], which also gives their eigenvectors and eigenvalues.

Mixing Matrices Involving $q\bar{q}$ Pairs

For $q\bar{q}$ and $gq\bar{q}$ in the bases used in Eq. (5.141) and Eqs. (5.110) and (5.149), we have

$$\widehat{\Delta}_{q\bar{q}}(\mu^2) = C_F\, L_{12}\,, \qquad \widehat{\Delta}_{g\,q\bar{q}}(\mu^2) = \frac{1}{2}\Big[C_A L_{12\cdot13} + (2C_F - C_A)L_{23}\Big]\,. \tag{5.263}$$

For $q\bar{q}q'\bar{q}'$ in the basis used in Eqs. (5.118), (5.186), and (5.167), we have

$$\widehat{\Delta}_{q\bar{q}\,q\bar{q}}(\mu^2) = \widehat{\Delta}_{q\bar{q}\,q'\bar{q}'}(\mu^2) = \begin{pmatrix} C_F\, L_{14\cdot23} + \left(C_F - \frac{1}{2}C_A\right) L_{12\cdot34/(13\cdot24)} & T_F\, L_{14\cdot23/(13\cdot24)} \\ T_F\, L_{12\cdot34/(13\cdot24)} & C_F\, L_{12\cdot34} + \left(C_F - \frac{1}{2}C_A\right) L_{14\cdot23/(13\cdot24)} \end{pmatrix}. \tag{5.264}$$

For $ggq\bar{q}$ in the basis used in Eqs. (5.124), (5.192), and (5.177), we have

$$\widehat{\Delta}_{gg\,q\bar{q}}(\mu^2) = \begin{pmatrix} \frac{1}{2}C_A\, L_{12\cdot13\cdot24} + \left(C_F - \frac{1}{2}C_A\right) L_{34} & 0 & T_F\, L_{13\cdot24/(14\cdot23)} \\ 0 & \frac{1}{2}C_A\, L_{12\cdot14\cdot23} + \left(C_F - \frac{1}{2}C_A\right) L_{34} & T_F\, L_{14\cdot23/(13\cdot24)} \\ T_F\, L_{12\cdot34/(14\cdot23)} & T_F\, L_{12\cdot34/(13\cdot24)} & C_A\, L_{12} + C_F\, L_{34} \end{pmatrix}. \tag{5.265}$$

Again, these simplify in the dijet case, for which they were given along with their eigenvectors and eigenvalues in Ref. [335].

5.8.4 Soft Function Evolution

In this section, we review the renormalization group evolution of the soft function, focusing on our use of the color basis notation of Sec. 5.3.4 for nonorthogonal bases. We will consider

308

the particular case of the N-jettiness event shape [104], which allows for a definition of exclusive N-jet production with a factorization theorem of the form of Eq. (5.2).

The color mixing matrices of the previous section are in general complex valued for physical kinematics. For a physical channel, some of the appearing s_{ij} are positive, giving rise to imaginary terms from the logarithms, as in Eq. (5.73). Since the cross section is real, these imaginary terms generated by the renormalization group evolution must drop out of the final result. We start by describing the properties of the soft function that ensure that this is the case.

Recall that the hard function $\widehat{H}_\kappa$ for a particular partonic channel κ has its color indices contracted with those of the soft function. Explicitly,

$$\text{tr}(\widehat{H}_\kappa \widehat{S}_\kappa) = H_\kappa^{a_1 \cdots a_n b_1 \cdots \beta_n} S_\kappa^{b_1 \cdots \beta_n a_1 \cdots \alpha_n} \tag{5.266}$$
$$= \sum_{\{\lambda_i\}} \left[C_{\lambda_1 \cdots (\cdot \lambda_n)}^{b_1 \cdots \beta_n} \right]^* S_\kappa^{b_1 \cdots \beta_n a_1 \cdots \alpha_n} C_{\lambda_1 \cdots (\cdot \lambda_n)}^{a_1 \cdots \alpha_n}.$$

The soft function is defined as a vacuum matrix element of a product of soft Wilson lines $\widehat{Y}$ as

$$\widehat{S}_\kappa(M, \{n_i\}) = \left\langle 0 \middle| \bar{\text{T}} \, \widehat{Y}^\dagger(\{n_i\}) \, \delta(M - \hat{M}) \, \text{T} \, \widehat{Y}(\{n_i\}) \middle| 0 \right\rangle, \tag{5.267}$$

where $\widehat{Y}(\{n_i\})$ is a product of soft Wilson lines in the n_i directions. It is a matrix in color space, and $\widehat{Y}^\dagger$ is its Hermitian conjugate. Here T and $\bar{\text{T}}$ denote time ordering and antitime ordering respectively. The matrices $\widehat{Y}$ and $\widehat{Y}^\dagger$ are multiplied with each other, i.e. one of the color indices of the corresponding Wilson lines are contracted, and the external indices correspond to $b_1 \cdots \beta_n$ and $a_1 \cdots \alpha_n$, respectively. Thus, for example $\widehat{Y}^\dagger \widehat{Y} = \delta^{a_1 b_1} \cdots \delta^{\alpha_n \beta_n}$. The dependence of the soft function on the particular measurement, as well as the details of the jet algorithm, are encoded in the measurement function $\hat{M}$, whose precise form is not relevant for the current discussion.

From the definition of the soft function in Eq. (5.267) we see that it is Hermitian, namely $(S_\kappa^{b_1 \cdots \beta_n a_1 \cdots \alpha_n})^* = S_\kappa^{a_1 \cdots \alpha_n b_1 \cdots \beta_n}$. In abstract notation, this means $\widehat{S}_\kappa^\dagger = \widehat{S}_\kappa$, which implies that

309

the product $\vec{C}^\dagger \widehat{S}_\kappa \vec{C}$ appearing in the cross section is real, so imaginary terms that appear in the Wilson coefficients due to renormalization group evolution drop out in the final cross section.

While this argument is trivial in a basis independent form, it is important to emphasize that in a nonorthogonal basis it takes a slightly more complicated form. As discussed in Sec. 5.3.4, in a specific nonorthogonal color basis, Eq. (5.266) takes the form $\vec{C}^\dagger \widehat{S}_\kappa \vec{C} = \vec{C}^{*T} \widehat{T} \widehat{S}_\kappa \vec{C}$ as in Eq. (5.58), where the matrix $\widehat{T}$ is defined in Eq. (5.56). Similarly, the matrix representation of $\widehat{S}_\kappa$ is not Hermitian with respect to the naive conjugate transpose of its components. Instead, the condition on the reality of the cross section is given by [see Eq. (5.57)]

$$\widehat{S}_\kappa = \widehat{S}_\kappa^\dagger = \widehat{T}^{-1} \widehat{S}_\kappa^{*T} \widehat{T} \,. \tag{5.268}$$

The invariance of the cross section under the RGE

$$\mu \frac{\mathrm{d}}{\mathrm{d}\mu} \sigma_N = 0 \,, \tag{5.269}$$

implies relations between the anomalous dimensions of the SCET functions appearing in the factorization theorem of Eq. (5.2). In particular, it allows the anomalous dimension of the soft function to be determined from the anomalous dimensions of the Wilson coefficients, along with the anomalous dimensions of the beam and jet functions. The anomalous dimensions of the jet and beam functions are proportional to the color-space identity. The anomalous dimensions of the beam and jet functions appearing in the N-jettiness factorization theorem are equal to all orders in perturbation theory [392] allowing us to use only the jet function anomalous dimension in the following discussion. Renormalization group consistency then implies that the contributions of the soft function anomalous dimension not proportional to the identity, including the color off-diagonal components, are completely determined by the anomalous dimensions of the Wilson coefficients.

The soft function for N-jettiness can be written in the general form of Eq. (5.267), but

with an explicit measurement function

$$\widehat{S}_\kappa(k_a, k_b, k_1, \ldots, k_N, \{n_i\}) \tag{5.270}$$

$$= \Big\langle 0 \Big| \bar{\mathrm{T}} \, \widehat{Y}^\dagger(\{n_i\}) \prod_i \delta(k_i - \hat{\mathcal{T}}_i) \, \mathrm{T} \, \widehat{Y}(\{n_i\}) \Big| 0 \Big\rangle .$$

Here $\hat{\mathcal{T}}_i$ picks out the contribution to the N-jettiness observable from the momentum region i, whose precise definition can be found in Ref. [301]. The soft function for N-jettiness was first presented to NLO in Ref. [301], and more recently analyzed to NNLO in Ref. [211].

The all-orders structure of the renormalization group evolution for the soft function can be derived from Eq. (5.269), and is given by [296, 301]

$$\mu \frac{\mathrm{d}}{\mathrm{d}\mu} \widehat{S}_\kappa(\{k_i\}, \mu)$$

$$= \int \Big[\prod_i \mathrm{d}k_i' \Big] \frac{1}{2} \Big[\widehat{\gamma}_S(\{k_i - k_i'\}, \mu) \, \widehat{S}_\kappa(\{k_i'\}, \mu)$$

$$+ \widehat{S}_\kappa(\{k_i'\}, \mu) \, \widehat{\gamma}_S^\dagger(\{k_i - k_i'\}, \mu) \Big] . \tag{5.271}$$

The soft anomalous dimension $\widehat{\gamma}_S$, and its conjugate $\widehat{\gamma}_S^\dagger$, are given in terms of the anomalous dimension γ_J of the jet function and the anomalous dimension of the Wilson coefficients, $\widehat{\gamma}_C$ defined in Eqs. (5.230) and (5.231), as

$$\widehat{\gamma}_S(\{k_i\}, \mu) = -\mathbf{1} \sum_i Q_i \, \gamma_J^i(Q_i k_i, \mu) \prod_{j \neq i} \delta(k_j)$$

$$- 2\widehat{\gamma}_C^\dagger(\mu) \prod_i \delta(k_i) . \tag{5.272}$$

(Here, the Q_i are related to the precise N-jettiness definition, see Ref. [301].) The Hermitian conjugates of $\widehat{\gamma}_C$ and $\widehat{\gamma}_S$ above again refer to the abstract Hermitian conjugate in color space. In a nonorthogonal color basis, they are given in terms of the complex conjugate transpose

311

components according to Eq. (5.57) as

$$\widehat{\gamma}_C^\dagger = \widehat{T}^{-1} \widehat{\gamma}_C^{*T} \widehat{T}, \qquad \widehat{\gamma}_S^\dagger = \widehat{T}^{-1} \widehat{\gamma}_S^{*T} \widehat{T}. \qquad (5.273)$$

5.9 Conclusions

In this chapter, we have presented a helicity operator approach to SCET. Helicities are naturally defined with respect to the external lightlike reference vectors specifying the jet directions in the effective theory, eliminating the need to consider complicated Lorentz and gamma matrix structures in the operator basis. The helicity operators correspond directly to physical states of definite helicity and color, which when combined with color organization techniques, greatly simplifies the construction of a minimal operator basis. Furthermore, the helicity operators are automatically crossing symmetric, and make manifest parity and charge conjugation symmetries, making it simple to determine relations amongst Wilson coefficients.

We demonstrated the utility of the helicity operator approach by explicitly constructing the basis valid to all orders in perturbation theory for a number of key processes at the LHC involving jets, and then determining the matching coefficients. In particular we considered $pp \to H + 0, 1$ jets, $pp \to W/Z/\gamma + 0, 1$ jets, and $pp \to 2$ jets at next-to-leading order, and $pp \to H + 2$ jets, $pp \to W/Z/\gamma + 2$ jets, and $pp \to 3$ jets at leading order. We also discussed the dependence of this matching on the regularization scheme, considering schemes with helicities in 4 and d dimensions. An important and well-known simplification of the SCET approach is that when dimensional regularization is used for both IR and UV divergences, all loop graphs in the effective theory are scaleless, and thus vanish. As a result, the hard SCET Wilson coefficients in the $\overline{\text{MS}}$ scheme, determined from matching QCD to SCET, are given directly by the IR-finite parts of color-ordered helicity amplitudes, defined using Eq. (5.96). The use of our helicity operator basis therefore makes it simple to combine analytic resummation in SCET with fixed-order calculations of helicity amplitudes.

The all-orders structure for the renormalization group evolution of the helicity operator

basis was discussed in detail. At leading power, distinct helicity structures do not mix, with renormalization group evolution causing mixing only in color space. This feature is made manifest at the level of the SCET Lagrangian due to the expansion in the soft and collinear limits. Subtleties associated with the use of nonorthogonal color bases were carefully treated, and expressions for the color sum matrix $\widehat{T}$ are given for the used color bases for all processes considered in the chapter. Explicit results are also given for the one-loop mixing matrices describing the renormalization group evolution in color space for the case of $pp \rightarrow$ up to 2 jets with an arbitrary number of uncolored external particles and in a manifestly crossing symmetric form.

Combining the methods of this chapter with known expressions for jet, beam, and soft functions for particular exclusive jet cross sections, or jet shapes/observables, should facilitate analytic resummation for a large number of processes for which fixed-order amplitudes are known, or are soon to be calculated.

Chapter 6

Jet Vetoes Interfering with $H \to WW$

Far off-shell Higgs production in $H \to WW, ZZ$, is a particularly powerful probe of Higgs properties, allowing one to disentangle Higgs width and coupling information unavailable in on-shell rate measurements. These measurements require an understanding of the cross section in the far off-shell region in the presence of realistic experimental cuts. We analytically study the effect of a p_T jet veto on far off-shell cross sections, including signal-background interference, by utilizing hard functions in the soft collinear effective theory that are differential in the decay products of the W/Z. Summing large logarithms of $\sqrt{\hat{s}}/p_T^{veto}$, we find that the jet veto induces a strong dependence on the partonic centre of mass energy, $\sqrt{\hat{s}}$, and modifies distributions in $\sqrt{\hat{s}}$ or M_T. The example of $gg \to H \to WW$ is used to demonstrate these effects at next to leading log order. We also discuss the importance of jet vetoes and jet binning for the recent program to extract Higgs couplings and widths from far off-shell cross sections.

6.1 Introduction

With the recent discovery of a boson resembling a 126 GeV Standard Model (SM) Higgs [393, 394], a large program has begun to study in detail the properties of the observed particle [395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410,

411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423]. Of fundamental interest are the couplings to SM particles and the total width of the observed boson, which is a sensitive probe of BSM physics [424, 425, 426, 427, 428, 429, 430]. Most studies have focused on the extraction of Higgs properties from on-shell cross sections. In this case, the effect of jet vetoes and jet binning, which is required experimentally in many channels to reduce backgrounds, has been well studied theoretically [300, 313, 315, 314, 208, 191, 431, 316, 13]. A jet veto, typically defined by requiring that there are no jets with $p_T \geq p_T^{veto}$, introduces large logarithms, $\log(m_H/p_T^{veto})$, potentially invalidating the perturbative expansion, and requiring resummation for precise theoretical predictions. In this chapter, we analytically study the effect of an exclusive jet p_T-veto on off-shell particle production, resumming logarithms of $\sqrt{\hat{s}}/p_T^{veto}$, where $\sqrt{\hat{s}}$ is the invariant mass of the off-shell particle. We use $gg \to H \to WW$ as an example to demonstrate these effects, although the formalism applies similarly to $gg \to H \to ZZ$ if a jet veto is imposed. We find that the off-shell cross section is significantly suppressed by a jet veto, and that the suppression has a strong dependence on $\sqrt{\hat{s}}$. This results in a modification of differential distributions in $\sqrt{\hat{s}}$, or any transverse mass variable, in the case that the invariant mass cannot be fully reconstructed. The jet veto also has an interesting interplay with signal-background interference effects, which typically contribute over a large range of $\sqrt{\hat{s}}$. We use two cases, $m_H = 126$ GeV, and $m_H = 600$ GeV, to demonstrate the effect of the jet veto on the signal-background interference in $gg \to H \to WW$.

There exist multiple motivations why it is important to have a thorough understanding of the far off-shell region in Higgs production, and the impact of a jet p_T veto on this region. As has been emphasized in a number of recent chapters [404, 432, 45, 433, 434], the separate extraction of the Higgs couplings and total width is not possible using only rate measurements for which the narrow width approximation (NWA) applies. In the NWA the cross section depends on the couplings and widths in the form

$$\sigma^{\text{nwa}} \sim \frac{g_i^2 g_f^2}{\Gamma_H}, \tag{6.1}$$

which is invariant under the rescaling

$$g_i \to \xi g_i, \qquad \Gamma_H \to \xi^4 \Gamma_H, \tag{6.2}$$

preventing their individual extraction from rate measurements alone.

The direct measurement of the width of the observed Higgs-like particle, expected to be close to its SM value of $\simeq 4\text{MeV}$, is difficult at the LHC, but is of fundamental interest as a window to new physics [424, 425, 426, 427, 428, 429, 430]. It is also important for model independent measurements of the Higgs couplings. Proposals to measure the Higgs width include those that rely on assumptions on the nature of electroweak symmetry breaking [404], direct searches for invisible Higgs decays [435, 436, 437, 438, 402], and a proposed measurement of the mass shift in $H \to \gamma\gamma$ relative to $H \to ZZ \to 4l$ caused by interference [432].

More recently, it has been proposed [45, 433, 434] that the Higgs width can be bounded by considering the far off-shell production of the Higgs in decays to massive vector bosons. In this region there is a contribution from signal-background interference [439, 440, 441, 442], and from far off-shell Higgs production [443, 444]. Far off-shell, the Higgs propagator is independent of Γ_H, giving rise to contributions to the total cross section that scale as

$$\sigma^{\text{int}} \sim g_i g_f, \qquad \sigma_H^{\text{off-shell}} \sim g_i^2 g_f^2, \tag{6.3}$$

for the signal-background interference and off-shell cross section respectively. The method proposed in [45] takes advantage of the fact that these components of the cross section scale differently than the NWA cross section. For example, in a scenario with large new physics contributions to the Higgs width, on-shell rate measurements at the LHC consistent with SM predictions enforce through Eq. (6.2) that the Higgs couplings are also scaled as $g_i \to g_i \left(\Gamma_H / \Gamma_H^{SM}\right)^{1/4}$. The off-shell and interference contributions to the cross section are

not invariant under this rescaling of the couplings, under which they are modified to

$$\sigma^{\text{int}} = \sqrt{\frac{\Gamma_H}{\Gamma_H^{SM}}}\sigma_{SM}^{\text{int}}, \qquad \sigma_H^{\text{off-shell}} = \frac{\Gamma_H}{\Gamma_H^{SM}}\sigma_{H,SM}^{\text{off-shell}}. \tag{6.4}$$

A measurement of the off-shell and interference cross section then allows for one to directly measure, or bound, the total Higgs width. This method is not completely model independent, indeed some of its limitations were recently discussed in [445], along with a specific new physics model which decorrelated the on-shell and off-shell cross sections, evading the technique. However, interpreted correctly, this technique places restrictions on the Higgs width in many models of BSM physics. The study of the off-shell cross section as a means to bound the Higgs width was first discussed in the $H \to ZZ \to 4l$ channel [45, 433], where the ability to fully reconstruct the invariant mass of the decay products allows for an easy separation of the on-shell and off-shell contributions. Recently, CMS has performed a measurement following this strategy and obtained a bound of $\Gamma_H \leq 4.2\ \Gamma_H^{SM}$ [446] .

The method was extended in [434] to the $gg \to H \to WW \to l\nu l\nu$ channel. The WW channel has the advantage that the $2W$ threshold is closer than for $H \to ZZ$, as well as having a higher branching ratio to leptons, and a higher total cross section. It does however, also have the disadvantage of large backgrounds, which necessitate the use of jet vetoes, as well as final state neutrinos which prevent the reconstruction of the invariant mass. To get around the latter issue one can exploit the transverse mass variable

$$M_T^2 = (E_T^{\text{miss}} + E_T^{ll})^2 - |\mathbf{p_T^{ll}} + \mathbf{E_T^{miss}}|^2, \tag{6.5}$$

which has a kinematic edge at $M_T = m_H$ for the signal. This variable was shown to be effective in separating the region where the off-shell and interference terms are sizeable, namely the high M_T region, from the low M_T region where on-shell production dominates, allowing for the extraction of a bound on the total Higgs width. Although the experimental uncertainties are currently large in the high M_T region, the authors estimate that with a reduction in the background uncertainty to $\lesssim 10\%$, the WW channel could be used to place

318

a bound on the Higgs width competitive with, and complementary to the bound from the $H \to ZZ \to 4l$ channel. They therefore suggest a full experimental analysis focusing on the high-M_T region of the WW channel. More generally, it was proposed in [447] that a similar method can also be used to probe couplings to heavy beyond the Standard Model states.

Independent of bounding the Higgs width, the study of the off-shell cross section opens up a new way to probe Higgs properties, which is particularly interesting as it probes particles coupling to the Higgs through loops over a large range of energies. Further benefits of the measurement of the off-shell cross section for constraining the parity properties of the Higgs, as well as for bounding higher dimensional operators were also discussed in [445].

A full theoretical understanding of the far off-shell region, especially in the presence of realistic experimental cuts, is therefore well motivated to allow for a proper theoretical interpretation of the data, and of bounds on new physics. Indeed, the current limits on the Higgs width from the off-shell region are based on leading order calculations combined with a parton shower. There has recently been progress on the calculation of the perturbative amplitudes required for an NLO description of the off-shell cross section, including signal-background interference, with the calculation of the two loop master integrals with off-shell vector bosons [448, 449]. However, one aspect that has not yet been addressed theoretically is the effect of jet vetoes, and more generally jet binning, on far off-shell cross sections, and on the signal-background interference.

Jet vetoes and jet binning are used ubiquitously in LHC searches to reduce backgrounds. They are typically defined by constraining the p_T of jets in the event. The $H \to WW$ channel is an example of such a search, where the exclusive zero jet bin, defined by enforcing that all jets in the event satisfy $p_T < p_T^{veto}$, is used to reduce the large background from $t\bar{t}$ production. Indeed, the analysis of [434] used the exclusive $N_{jet} = 0$ bin in the large M_T region to estimate the bound on the Higgs width achievable in the $H \to WW$ channel. Furthermore, the recent bound by CMS [446] of the Higgs width from the $H \to ZZ \to 2l2\nu$ channel used jet bins to optimize sensitivity, splitting data into exclusive 0-jet and inclusive 1-jet samples, which were each analyzed and then combined to give the limit. The proper

319

interpretation of the off-shell cross section measurements requires understanding, preferably analytically, the impact of the jet veto and jet binning procedures.

As is well known, the jet veto introduces a low scale, typically $p_T^{veto} \sim 25 - 30\text{GeV}$, into a problem which is otherwise characterized by the scale Q, of the hard collision. This causes large logarithms of the form $\alpha_s^n \log^m (Q/p_T^{veto})$, $m \leq 2n$, to appear in perturbation theory, forcing a reorganization of the perturbative expansion. Physically these logarithms arise due to constraints placed on the radiation in the event, which prevent a complete cancellation of real and virtual infrared contributions. A resummation to all orders in α_s is then required to make precise predictions. For the leading logarithms this resummation can be implemented by a parton shower. This approach is however difficult to systematically improve, and does not allow for higher order control of the logarithmic accuracy, or a systematic analysis of theoretical uncertainties in the correlations between jet bins. An alternative approach, which allows for the analytic resummation of large logarithms appearing in the cross section, is to match to the soft collinear effective theory (SCET) [21, 22, 23, 24, 25], which provides an effective field theory description of the soft and collinear limits of QCD. In SCET, large logarithms can be resummed through renormalization group evolution to desired accuracy, providing analytic control over the resummation. This framework also provides control over the theoretical uncertainties, including the proper treatment of correlations between jet bins [300, 450, 451].

The effect of jet vetoes on Higgs production in the on-shell region has attracted considerable theoretical interest [300, 313, 315, 314, 208, 191, 316, 13]. For on-shell Higgs production, $Q \sim m_H$, and hence the resummation is of logarithms of the ratio m_H/p_T^{veto}. The use of a jet clustering algorithm in the experimental analyses complicates resummation and factorization [167, 191], and leads to logarithms of the jet radius parameter [191, 312, 452, 315]. Current state of the art calculations achieve an NNLL'+NNLO accuracy [316, 13], along with the incorporation of the leading dependence on the jet radius, allowing for precise theoretical predictions in the presence of a jet veto. Such predictions are necessary for the reliable extractions of Higgs couplings from rate measurements. Indeed, the exclusive zero-jet Higgs

cross section is found to decrease sharply as the p_T^{veto} scale is lowered.

In this chapter we use SCET to analytically study the effect of a jet veto on off-shell cross sections. In particular, we are interested in processes with contributions from a large range of $\hat{s}$, where $\sqrt{\hat{s}}$ is the partonic centre of mass energy. In Sec. 6.2, we present a factorization theorem allowing for the resummation of large logarithms of the form $\log \sqrt{\hat{s}}/p_T^{veto}$, in the cross section for the production of a non-hadronic final state. Working to NLL order, and using canonical scales, for simplicity, gives [313]

$$\frac{d\sigma_0^{\mathrm{NLL}}(p_T^{veto})}{d\hat{s}\,d\Phi} = \left| \mathcal{M}_{ij}(\mu = \sqrt{\hat{s}}, \Phi) \right|^2 \int dx_a dx_b f_i(x_a, \mu = p_T^{veto}) f_j(x_b, \mu = p_T^{veto}) \tag{6.6}$$

$$\times\, \delta(x_a x_b E_{\mathrm{cm}}^2 - \hat{s}) e^{-2\mathrm{Re}K_{\mathrm{NLL}}^i(\sqrt{\hat{s}}, p_T^{veto})} \, ,$$

where $\sigma_0(p_T^{veto})$ is the exclusive zero-jet cross section. In this formula, f_i, f_j are the parton distribution functions (PDFs) for species i, j, $\mathcal{M}_{ij}$ is the hard matrix element, Φ is the leptonic phase space, and E_{cm} is the hadronic centre of mass energy. K_{NLL}^i is a Sudakov factor, defined explicitly in Sec. 6.2, which depends only on the identity of the incoming partons. The form of Eq. (6.6) shows that the effect of the jet veto can be captured independent of the hard underlying process, which enters into Eq. (6.6) only through $\mathcal{M}$. At higher logarithmic order a dependence on the jet algorithm is also introduced, but the ability to separate the effect of the jet veto from the particular hard matrix element using the techniques of factorization remains true, and allows one to make general statements about the effect of the jet veto.

The resummation of the large logarithms, $\log \sqrt{\hat{s}}/p_T^{veto}$, introduced by the jet veto leads to a suppression of the exclusive zero-jet cross section, evident in Eq. (6.6) through the Sudakov factor, and familiar from the case of on-shell production. The interesting feature in the case of off-shell effects is that this suppression depends on $\sqrt{\hat{s}}$. For example, when considering off-shell Higgs production, or signal-background interference, which contribute over a large range of $\sqrt{\hat{s}}$, the jet veto re-weights contributions from different $\sqrt{\hat{s}}$ regions in a strongly $\sqrt{\hat{s}}$ dependent manner. In particular, this modifies differential distributions in $\sqrt{\hat{s}}$, or any

321

similar variable, such as M_T. This is of particular interest for the program to place bounds on the Higgs width using the off-shell cross section in channels which require a jet veto, as this procedure requires an accurate description of the shape of the differential cross section. Furthermore, the jet veto has an interesting effect on the signal-background interference, which often exhibits cancellations from regions widely separated in $\sqrt{\hat{s}}$. The study of these effects is the subject of this chapter.

Our outline is as follows. In Sec. 6.2 we review the factorization theorem for the exclusive zero jet bin, with a jet veto on the p_T of anti-k_T jets, focussing on the dependence on $\sqrt{\hat{s}}$. Sec. 6.3 describes the generic effects of jet vetoes on off-shell production, including the dependence on the jet veto scale, the identity of the initial state partons and the hadronic centre of mass energy. In particular, we show that off-shell production in the exclusive zero-jet bin is suppressed by a strongly $\sqrt{\hat{s}}$ dependent Sudakov factor, and comment on the corresponding enhancement of the inclusive 1-jet cross section. In Sec. 6.4 we perform a case study for the $gg \to H \to WW \to l\nu l\nu$ process, resumming to NLL accuracy the off-shell cross section including the signal-background interference. For the signal-background interference, we consider two Higgs masses, $m_H = 125$ GeV and $m_H = 600$ GeV, whose interference depends differently on $\sqrt{\hat{s}}$, to demonstrate different possible effects of the jet veto on the signal-background interference. Since $\sqrt{\hat{s}}$ is not experimentally reconstructible for $H \to WW$, in Sec. 6.4.5, we demonstrate the suppression as a function of M_T caused by the jet-veto restriction. In Sec. 6.5 we discuss the effect of the jet veto and jet binning on the extraction of the Higgs width from the off-shell cross section in $H \to WW$ (commenting also on $H \to ZZ$). We conclude in Sec. 6.6.

6.2   Cross Sections with a Jet Veto: A Review

In this section we review the factorization theorem, in the SCET formalism, for $pp \to L + 0$-jets, where L is a non-hadronic final state. We consider a jet veto defined by clustering an event using an anti-k_T algorithm with jet radius R to define jets, J, and imposing the constraint that $p_T^J < p_T^{veto}$ for all jets in the event. This is the definition of the jet veto

currently used in experimental analyses, with the experimental value of p_T^{veto} typically $25-30$ GeV, and $R \simeq 0.4$-0.5.

6.2.1 Factorization Theorem

Following the notation of [13], the factorization theorem for $pp \to L + 0$-jets with a jet veto on p_T can be computed in the framework of SCET. For a hard process where L has invariant mass $\sqrt{\hat{s}}$ (on-shell or off-shell), we have

$$
\frac{d\sigma_0(p_T^{veto})}{d\hat{s}} = \int d\Phi dx_a dx_b \; \delta(x_a x_b E_{cm}^2 - \hat{s}) \sum_{i,j} H_{ij}(\sqrt{\hat{s}}, \Phi, \mu) B_i(\sqrt{\hat{s}}, p_T^{veto}, R, x_a, \mu, \nu)
$$
$$
\times B_j(\sqrt{\hat{s}}, p_T^{veto}, R, x_b, \mu, \nu) S_{ij}(p_T^{veto}, R, \mu, \nu)
$$
$$
+ \frac{d\sigma_0^{Rsub}(p_T^{veto}, R)}{d\hat{s}} + \frac{d\sigma_0^{ns}(p_T^{veto}, R, \mu_{ns})}{d\hat{s}}. \tag{6.7}
$$

In this formula, Φ denotes the leptonic phase space, i, j denote the initial partonic species, H_{ij} is the hard function for a given partonic channel, B_i are the beam functions which contain the PDFs, and S_{ij} is the soft function, each of which will be reviewed shortly. Since this factorization theorem applies to the production of a color singlet final state, we either have $i = j = g$, or $i = q$, $j = \bar{q}$. Eq. (6.7) is written as the sum of three terms. The first term in Eq. (6.7) contains the singular logarithmic terms, which dominate as $p_T^{veto} \to 0$, or in the case of off-shell production that we are considering, as $\hat{s} \to \infty$, with p_T^{veto} fixed. The second term, σ_0^{Rsub}, contains corrections that are polynomial in the jet radius parameter R, and σ_0^{ns} contains non-singular terms which vanish as $p_T^{veto} \to 0$, and are suppressed relative to the singular terms when the ratio $p_T^{veto}/\sqrt{\hat{s}}$ is small.

The factorization theorem allows for each component of Eq. (6.7) to be calculated at its natural scale, and evolved via renormalization group evolution (RGE) to a common scale, resumming the large logarithms of $p_T^{veto}/\sqrt{\hat{s}}$. For the case of a veto on the jet p_T, the factorization follows from SCET$_{\mathrm{II}}$, where the RGE is in both the virtuality scale, μ, and rapidity scale, ν [453, 308]. In this section, we will briefly summarize the components of the factorization theorem with a particular focus on their dependence on the underlying hard

matrix element, the identity of the incoming partons, the jet algorithm, and the jet veto measurement. We will also review their RGE properties. Further details, including analytic expressions for the anomalous dimensions, can be found in [191, 13, 300], and references therein.

Soft Function

The soft function $S_{ij}(p_T^{veto}, R, \mu, \nu)$ describes the soft radiation from the incoming partons i, j which are either both gluons or both quarks. It is defined as a matrix element of soft Wilson lines along the beam directions, with a measurement operator, $\mathcal{M}^{jet}$, which enforces the jet veto condition:

$$S_{ii}(p_T^{veto}, R, \mu, \nu) = \langle 0| Y_{n_b} Y_{n_a}^\dagger \mathcal{M}^{jet}(p_T^{veto}, R) Y_{n_a} Y_{n_b}^\dagger |0\rangle . \qquad (6.8)$$

The soft function depends only on the identity of the incoming partons, through the representation of the Wilson lines, which has not been made explicit. It also depends on the definition of the jet veto through the measurement function.

The soft function is naturally evaluated at the soft scale $\mu_S \sim p_T^{veto}$, and $\nu_S \sim p_T^{veto}$, and satisfies a multiplicative renormalization in both μ and ν. The solution is given by

$$S_{ii}(p_T^{veto}, R, \mu, \nu) = S_{ii}(p_T^{veto}, R, \mu_S, \nu_S) \exp\left[\log\frac{\nu}{\nu_S}\gamma_\nu^i(p_T^{veto}, R, \mu_S)\right] \exp\left[\int_{\mu_S}^{\mu} \frac{d\mu'}{\mu'}\gamma_S^i(\mu', \nu)\right] .$$

$$(6.9)$$

Further details including expressions for the anomalous dimensions are given in [13].

In the case of interest, where the jets are defined using a clustering algorithm with a finite R, the soft function also contains clustering logarithms from the clustering of correlated soft emissions, which first arise at NNLL. These appear in the cross section as logarithms of the jet radius, $\log(R)$, but are not resummed by the RGE. For experimentally used values of R, the first of these logarithms is large [315], while the leading $\mathcal{O}(\alpha_s^3)$ term was recently

calculated and found to be small [452]. We therefore treat these $\log(R)$ factors in fixed order perturbation theory. We discuss the impact of these logarithms on our results in Sec. 6.4.3.

Beam Function

The beam function [454, 20, 392], B_i, describes the collinear initial state-radiation from an incoming parton, i, as well as its extraction from the colliding protons through a parton distribution function. The beam function depends only on the identity of the incoming parton i, and the measurement function.

In the case of a p_T^{veto}, the beam function can be calculated perturbatively by matching onto the standard PDFs at the beam scale $\mu_B \sim p_T^{veto}$, $\nu_B \sim \sqrt{\hat{s}}$:

$$B_i(\sqrt{\hat{s}}, p_T^{veto}, R, x, \mu_B, \nu_B) = \sum_j \int_x^1 \frac{dz}{z} \mathcal{I}_{ij}(\sqrt{\hat{s}}, p_T^{veto}, R, z, \mu_B, \nu_B) f_j\left(\frac{x}{z}, \mu_B\right) \tag{6.10}$$

The lowest order matching coefficient is

$$\mathcal{I}_{ij} = \delta_{ij}\delta(1-z) \tag{6.11}$$

so that to leading order the beam function is simply the corresponding PDF, but evaluated at the beam scale $\mu_B \simeq p_T^{veto}$. This was seen explicitly in the NLL expansion of Eq. (6.6). Higher order matching coefficients involve splitting functions, allowing for a mixing between quarks and gluons. This matching procedure corresponds to the measurement of the proton at the scale $\mu_B \sim p_T^{veto}$ by the jet veto. Above the scale μ_B, the beam function satisfies a multiplicative RGE in both virtuality, μ and rapidity, ν, describing the evolution of an incoming jet for the off-shell parton of species i. Unlike the RGE for the PDFs, the RGE for the beam function leaves the identity and momentum fraction of the parton unchanged. The solution to the RGE is given by

$$B_i(\sqrt{\hat{s}}, p_T^{veto}, R, x, \mu, \nu) = B_i(\sqrt{\hat{s}}, p_T^{veto}, R, x, \mu_B, \nu_B) \exp\left[\frac{1}{2}\log\frac{\nu_B}{\nu}\gamma_\nu^i(p_T^{veto}, R, \mu_B)\right]$$

325

$$\times \exp\left[\int_{\mu_B}^{\mu} \frac{d\mu'}{\mu'} \gamma_B^i(\sqrt{\hat{s}}, \mu', \nu) \right],\tag{6.12}$$

which resums the logarithmic series associated with the collinear radiation from the incoming parton. Further details and expressions for the anomalous dimensions are again given in [13].

As with the soft function, the beam function also contains logarithms and polynomial dependence on the jet radius, R, from the clustering of collinear emissions. These logarithms can be numerically significant, but are not resummed by the RGE. We again treat these terms in fixed order perturbation theory.

Hard Function

The hard function H_{ij} encodes the dependence of the singular term of Eq. (6.7) on the underlying hard partonic matrix element of the $pp \to L + 0$-jets process. It can be obtained by matching QCD onto an appropriate SCET operator at the scale $\sqrt{\hat{s}}$, giving a Wilson coefficient, C_{ij}. The Wilson coefficient satisfies a standard RGE in virtuality, allowing it to be evolved to the scale μ. The hard function is then given by the square of the Wilson coefficient

$$H_{ij}(Q, \mu) = |C_{ij}(Q, \mu)|^2,\tag{6.13}$$

where Q denotes dependence on all variables associated with the final leptons as well as parameters like the top-mass, and the Higgs and W/Z masses and widths. The solution to the RGE equation for the hard function is

$$H_{ij}(Q, \mu) = H_{ij}(Q, \mu_H) \left| e^{-K^i(\sqrt{\hat{s}}, \mu_H, \mu)} \right|^2,\tag{6.14}$$

where the Sudakov form factor is

$$K^i(\sqrt{\hat{s}}, \mu_H, \mu) = \int_{\mu_H}^{\mu} \frac{d\mu'}{\mu'} \gamma_H^i(\sqrt{\hat{s}}, \mu')$$

$$= 2K_{\Gamma_{\text{cusp}}^i}(\mu_H, \mu) - K_{\gamma_H^i}(\mu_H, \mu) - \ln\left(\frac{-\hat{s} - i0}{\mu_H^2}\right) \eta_{\Gamma_{\text{cusp}}^i}(\mu_H, \mu).\tag{6.15}$$

Here the integrals involve the β-function and anomalous dimensions

$$K_{\Gamma_{\text{cusp}}^i} = \int_{\alpha_s(\mu_H)}^{\alpha_s(\mu)} \frac{d\alpha_s}{\beta(\alpha_s)} \Gamma_{\text{cusp}}^i(\alpha_s) \int_{\alpha_s(\mu_H)}^{\alpha_s} \frac{d\alpha_s'}{\beta(\alpha_s')}, \qquad \eta_{\Gamma_{\text{cusp}}^i} = \int_{\alpha_s(\mu_H)}^{\alpha_s(\mu)} \frac{d\alpha_s}{\beta(\alpha_s)} \Gamma_{\text{cusp}}^i(\alpha_s),$$

$$K_{\gamma_H^i} = \int_{\alpha_s(\mu_H)}^{\alpha_s(\mu)} \frac{d\alpha_s}{\beta(\alpha_s)} \gamma_H^i(\alpha_s), \tag{6.16}$$

where the channel i is either for quarks or gluons. Here the cusp and regular anomalous dimensions are $\Gamma_{\text{cusp}}^i(\alpha_s) = \sum_{k=0}^{\infty} \Gamma_k^i (\alpha_s/4\pi)^{k+1}$, $\gamma_H^i(\alpha_s) = \sum_{k=0}^{\infty} \gamma_k^i (\alpha_s/4\pi)^{k+1}$, respectively, and the β-function is $\beta(\alpha_s) = -2\alpha_s \sum_{k=0}^{\infty} \beta_k (\alpha_s/4\pi)^{k+1}$ so $\beta_0 = 11 C_A/3 - 2n_f/3$.

Explicit results for the functions in Eq. (6.16) can be found for example in Ref. [300]. Since we will be considering far off-shell production, and including signal-background interference effects, which have not been discussed in SCET factorization theorems before, we will discuss in more detail the definition of the hard function for the specific case of $gg \to l\nu l\nu$ in Sec. 6.4.1.

The beam and soft functions are universal, depending only on the given measurement and the identity of the incoming partons, it is the hard function that needs to be calculated separately for different processes. The beam and soft functions are known to NNLL for the case of a jet veto defined using a cut on p_T, and it is the hard coefficient that prevents resummation to NNLL for several cases of interest. In particular, since we are interested here in the case of off-shell production, one needs the full top mass dependence of loops, significantly complicating the computation. Indeed, for the case of signal-background interference for $gg \to H \to WW \to l\nu l\nu$, only the leading order hard function is known [439], while for direct gluon-fusion Higgs production, analytic results exist for the NLO virtual corrections including quark mass dependence [349]. This restricts our predictions to NLL accuracy for signal-background interference for $gg \to H \to WW \to l\nu l\nu$.

Non-Singular Terms

The non-singular term $\sigma_0^{ns}(p_T^{veto}, R, \mu_{ns})$ is an additive correction to the factorization theorem, containing terms that vanish as $p_T^{veto} \to 0$. This term scales as $p_T^{veto}/\sqrt{\hat{s}}$. The non-singular

piece is important when p_T^{veto} is of the same order as $\sqrt{\hat{s}}$, where both singular and non-singular pieces contribute significantly to the cross section. In this chapter, we will be focusing on the effect of a jet veto on far off-shell effects, and we will therefore always be considering the case that $p_T^{veto} \ll \sqrt{\hat{s}}$. We will therefore not discuss the non-singular pieces of the cross section, and focus on the singular contributions.

Uncorrelated Emissions

Beginning with two emissions, the jet algorithm can cluster uncorrelated emissions from the soft and collinear sectors [191, 314, 315]. This produces terms proportional to powers of R^2, which can formally be treated as power corrections for $R \ll 1$, and are included in σ_0^{Rsub}. For the jet radii of 0.4-0.5 used by the experimental collaborations, these effects are numerically very small, especially compared to the $\log R$ terms from correlated emissions. We make use of the expressions from [13].

6.2.2 Expansion to NLL

It is useful to consider the factorization theorem at NLL order with canonical scale choices, to see the main factors that control its behaviour. The result at NLL was first given in [313] for on-shell production with $\sqrt{\hat{s}} = m_H$. Allowing for off-shell production, and using canonical scales, the cross section with a p_T^{veto} cut is given at NLL by

$$\frac{d\sigma_0^{\mathrm{NLL}}(p_T^{veto})}{d\hat{s}\,d\Phi} = \left|\mathcal{M}_{ij}(\mu = \sqrt{\hat{s}}, \Phi)\right|^2 \int dx_a dx_b f_i(x_a, \mu = p_T^{veto}) f_j(x_b, \mu = p_T^{veto}) \tag{6.17}$$
$$\times\, \delta(x_a x_b E_{\mathrm{cm}}^2 - \hat{s}) e^{-2\mathrm{Re}K_{\mathrm{NLL}}^i(\sqrt{\hat{s}}, p_T^{veto})},$$

where Φ are phase space variables for the final state leptonic decay products. In this equation, f_i and f_j are the appropriate PDFs, for example, they are both f_g for the case of gluon-fusion since direct contributions from the quark PDFs do not enter until NNLL order. For a partonic center of mass energy $\sqrt{\hat{s}}$, Eq. (6.17) resums to NLL accuracy the logarithms of $\sqrt{\hat{s}}/p_T^{veto}$. Eq. (6.17) does not include the non-singular contribution to the cross section. As discussed

previously, in the far off-shell region, $p_T^{veto} \ll \sqrt{\hat{s}}$, and the singular contributions to the cross section dominate. It should also be emphasized that at NLL one is not sensitive to the jet algorithm or jet radius, as at $\mathcal{O}(\alpha_s)$ there is only a single soft or collinear emission. Although the R dependence is important for accurate numerical predictions, it does not effect the qualitative behaviour of the jet veto. The R dependence appears in the factorization theorem at NNLL.

The only dependence on the hard partonic process in Eq. (6.17) is in the matrix element $\mathcal{M}_{ij}(\hat{s})$. The Sudakov form factor K^i given in Eq. (6.15) arises from restrictions on real radiation in QCD, and depends only on the identity of the incoming partons. At NLL the Sudakov factor is given by

$$
K_{\mathrm{NLL}}^i(\sqrt{\hat{s}},\mu_H,\mu) = -\frac{\Gamma_0^i}{2\beta_0^2}\left\{\frac{4\pi}{\alpha_s(\mu_H)}\left(1 - \frac{1}{r} - \ln r\right) + \left(\frac{\Gamma_1^i}{\Gamma_0^i} - \frac{\beta_1}{\beta_0}\right)(1 - r + \ln r) + \frac{\beta_1}{2\beta_0}\ln^2 r\right\}
$$
$$
+ \frac{\gamma_0^i}{2\beta_0}\ln r + \ln\left(\frac{-\hat{s} - i0}{\mu_H^2}\right)\frac{\Gamma_0^i}{2\beta_0}\left\{\ln r + \frac{\alpha_s(\mu_H)}{4\pi}\left(\frac{\Gamma_1^i}{\Gamma_0^i} - \frac{\beta_1}{\beta_0}\right)(r - 1)\right\},
$$

$$(6.18)$$

where $r = \alpha_s(\mu)/\alpha_s(\mu_H)$. The form in Eq. (6.18) allows for the use of complex scales, such as $\mu_H = -i\sqrt{\hat{s}}$, to minimize the appearance of large π^2 factors in the Hard function. On the other hand, with canonical scales we would take $K_{\mathrm{NLL}}^i = K_{\mathrm{NLL}}^i(\sqrt{\hat{s}}, p_T^{veto})$. At LL order the terms with Γ_1, β_1, and γ_0 do not yet contribute and using the LL running coupling we can write $\mathrm{Re}K_{\mathrm{LL}}^i(\sqrt{\hat{s}}, p_T^{veto}) = -(4C/\beta_0)\ln\sqrt{\hat{s}}/p_T^{veto}\,[1 + \ln(1 - 2\lambda)/(2\lambda)]$ where $\lambda = \alpha_s(\sqrt{\hat{s}})\frac{\beta_0}{4\pi}\ln\sqrt{\hat{s}}/p_T^{veto}$. For gluon-fusion, $C = C_A = 3$, whereas for a quark-antiquark initial state, $C = C_F = 4/3$.

There are two important features of the expression in Eq. (6.17) compared with the case of no jet veto. First, the PDFs are evaluated at the scale $\mu = p_T^{veto}$ instead of $\mu = \sqrt{\hat{s}}$. Secondly, the cross section is multiplied by a Sudakov factor, which depends on logs of the ratio $\sqrt{\hat{s}}/p_T^{veto}$. These have a strong impact on the cross section, which will be the focus of Sec. 6.3.

329

6.3 Jet Vetoes and Off-Shell Effects

In this section we will discuss quite generally the effect of jet vetoes on off-shell cross sections. We focus on the dependence on the identity of the initial state partons, and the relation between the exclusive 0-jet and inclusive 1-jet bins. We conclude with a discussion of the dependence on the hadronic centre of mass energy. For simplicity, in this section we will use the NLL expansion of Eq. (6.17) with canonical scale choices. The NLL expansion demonstrates the essential features that persist at higher logarithmic order, and makes transparent how these effects depend on various parameters of interest. This serves for the purpose of demonstrating the generic effects of jet vetoes, and their dependencies. In Sec. 6.4, we will perform a more detailed study for the specific case of $gg \to H \to WW$.

Unlike on-shell effects, which contribute to the cross section over a small region in $\sqrt{\hat{s}}$, of order the width, off-shell effects, including signal-background interference and off-shell production, typically contribute over a large range of values of $\sqrt{\hat{s}}$. In this case the $\sqrt{\hat{s}}$ dependence of the jet veto suppression can produce interesting effects. In particular, it modifies differential distribution in $\sqrt{\hat{s}}$, or any substitute such as M_T in cases where the full invariant mass cannot be reconstructed, such as $H \to WW$. Furthermore, for signal-background interference, the $\sqrt{\hat{s}}$ dependence of the jet veto suppression can cause an enhancement or suppression of the interference relative to the on-shell contribution to the cross section, or enhance/suppress interference contributions with different signs relative to one another.

With this motivation, we now study the $\sqrt{\hat{s}}$ dependence of the jet veto suppression to the exclusive zero jet cross section using the NLL expression of Sec. 6.2.2. The benefit of the factorized expression is that this discussion can be carried out essentially independent of the matrix element prefactor $|\mathcal{M}_{ij}|^2$. From Eq. (6.17), the NLL cross section is modified compared with the LO cross section, only by the evaluation of the PDFs at the jet veto scale, and by the Sudakov factor, which is a function of $\sqrt{\hat{s}}$. To study the suppression due to the jet veto as a function of $\sqrt{\hat{s}}$, we will therefore consider

$$E_0(\hat{s}) = \left(\frac{d\sigma_0^{\mathrm{NLL}}(p_T^{veto})}{d\sqrt{\hat{s}}} \right) \bigg/ \left(\frac{d\sigma}{d\sqrt{\hat{s}}} \right) , \qquad (6.19)$$

where $\sigma_0^{NLL}(p_T^{veto})$ is the NLL exclusive zero jet cross section. In Eq. (6.19), the cross section in the denominator is evaluated to LO, namely to the same order as the matrix element that appears in Eq. (6.17) for the NLL resummed cross section. When forming this combination, one could choose to evaluate the denominator at various orders, for example using the full NLO result calculated without the p_T^{veto}. Since NLO corrections are typically large, especially for gluon initiated processes, this would typically decrease the above ratio. However, we have in mind an application to processes, such as signal-background interference in $gg \to l\nu l\nu$, for which the NLO corrections are not yet known, so that current calculations are restricted to LO results. In this case, we can incorporate the effect of the jet veto at NLL using Eq. (6.17), and the ratio of Eq. (6.19) will characterize the effect of the resummation compared to previous calculations in the literature [439, 433, 434]. This approach also has the benefit that it can be done independent of the particular matrix element, as the NLO corrections are clearly process dependent. However, all of the general features described in this section persist to NNLL resummation, as will be demonstrated in Sec. 6.4. As was mentioned previously, at NLL one doesn't have sensitivity to the jet radius R. While this dependence is important for precise predictions, it does not dominate the behaviour of the jet vetoed cross section as a function of $\hat{s}$, or modify in any way the conclusions of this section.

For numerical calculations in this section we use the NLO PDF fit of Martin, Stirling, Thorne and Watt [455] with $\alpha_s(m_Z) = 0.12018$. Unless otherwise stated, we use a hadronic center of mass energy of $E_{cm} = 8$ TeV. In Sec. 6.3.3 we discuss the dependence on the E_{cm}, comparing behaviour at $8, 13$, and 100 TeV.

In Fig. 6-1a we demonstrate the effect of the jet veto for a gluon-gluon initial state, as a function of $\sqrt{\hat{s}}$.[1] We plot the ratio $E_0(\hat{s})/E_0(m_H^2)$, for $m_H = 126$ GeV. We have chosen to plot this particular ratio to focus on the $\hat{s}$ dependence, rather than the impact that the jet veto has on the on-shell Higgs production cross section which is given by $E_0(m_H^2)$. The ratio $E_0(\hat{s})/E_0(m_H^2)$ describes the impact of the jet-veto for off-shell effects relative to its impact for on-shell production. It will also be useful when discussing the impact on Higgs width

---

[1]Note that a similar effect was considered in [300] which performed resummation for gluon fusion Higgs production with a veto on the global beam thrust event shape, as a function of the Higgs mass.
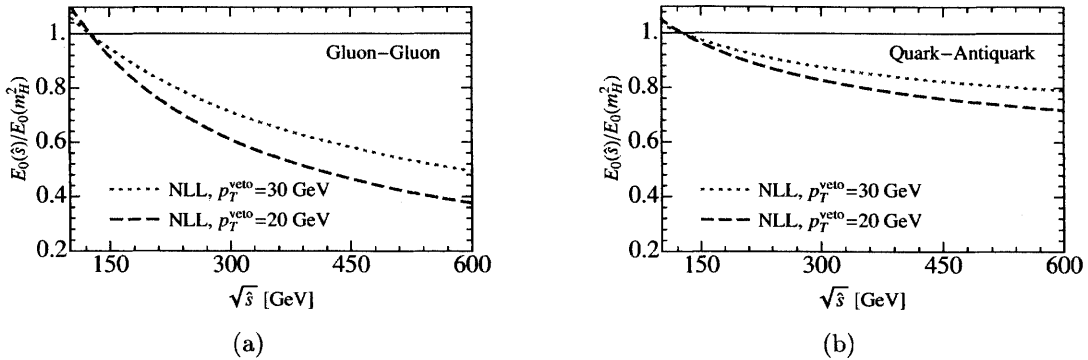
Figure 6-1: The ratio $E_0(\hat{s})/E_0(m_H^2)$, for both a gluon-gluon initiated process in (a), and a quark-antiquark initiated process in (b). In both cases we consider $p_T^{veto} = 20, 30$ GeV. The jet veto causes an $\hat{s}$ dependent suppression, which is significantly stronger for initial state gluons than initial state quarks, due to the larger colour factor appearing in the Sudakov.

bounds in Sec. 6.5. Fig. 6-1a shows that the suppression of the exclusive zero-jet cross section has a strong dependence on $\hat{s}$. The comparison between $p_T^{veto} = 20$ GeV, and $p_T^{veto} = 30$ GeV shows that a lower cut on the p_T of emissions causes a more rapid suppression, as expected. We have chosen to use the values $p_T^{veto} = 20$, 30 GeV, because CMS currently uses $p_T^{veto} = 30$ GeV, and although the ATLAS collaboration uses $p_T^{veto} = 25$ GeV, the $p_T^{veto} = 20$ GeV cut demonstrates the effects of a fairly extreme jet veto. Fig. 6-1a demonstrates that at scales of $\sqrt{\hat{s}} \simeq 500$ GeV, the suppression relative to that for on-shell production is of order 50%.

6.3.1 Quarks vs. Gluons

We now consider the difference in the jet veto suppression for quark initiated and gluon initiated processes. This is relevant in the case where multiple partonic channels contribute to a given process, or if the signal and background processes are predominantly from different partonic channels. This is the case for both $gg \to H \to WW$, ZZ, which have large $q\bar{q}$ initiated backgrounds. The factorization theorem in Eq. (6.7) allows one to easily study the dependence of the jet veto suppression on the identity of the incoming partons, which is carried by the hard, beam, and soft functions. The difference in the suppression arises from the differences in the anomalous dimensions, where for the 0-jet cross section, they

involve C_F for quarks, and C_A for gluons. The clustering and correlation logarithms are also multiplied by the colour factors C_F and C_A. This phenomenon is similar to quark vs. gluon discrimination for jets [48], where the same factors of C_F and C_A appear in the Sudakov and allow one to discriminate between quark and gluon jets. However, in this case, the discrimination is between incoming quarks and gluons.

Comparing Fig. 6-1a and Fig. 6-1b, we see a significant difference between a gluon-gluon and quark-antiquark initial state. The jet veto suppression increases more rapidly with $\hat{s}$ in the case of gluon-fusion induced processes than quark anti-quark induced processes. The suppression due to the jet veto being approximately twice as large for the case of gluon-fusion as for quark-antiquark fusion, for the values considered in Fig. 6-1. (Note that for the quark-antiquark initial state, we have used the up quark for concreteness, however, the result is approximately independent of flavour for the light quarks, with the suppression being dominated by the flavour independent Sudakov factor. A small dependence on flavour comes from the scale change in the PDF.) The effect of the jet veto is therefore of particular interest for gluon initiated processes, such as Higgs production through gluon-gluon fusion, to be discussed in Sec. 6.4. This difference in the suppression is interesting for a proper analysis of the backgrounds for $H \rightarrow WW$, ZZ in the off-shell region, and deserves further study since one may wish to vary p_T^{veto} as a function of $\sqrt{\hat{s}}$ or M_T.

6.3.2 Inclusive 1-Jet Cross Section

We have up to this point focused on the exclusive zero jet cross section. However, since the total inclusive cross section is unaffected by the jet veto, the inclusive 1-jet cross section has the same logarithmic structure as the exclusive zero-jet cross section, and can be related to the exclusive zero jet cross section by

$$\frac{d\sigma_{\geq 1}(p_T^{veto})}{d\sqrt{\hat{s}}} = \frac{d\sigma}{d\sqrt{\hat{s}}} - \frac{d\sigma_0(p_T^{veto})}{d\sqrt{\hat{s}}}. \tag{6.20}$$
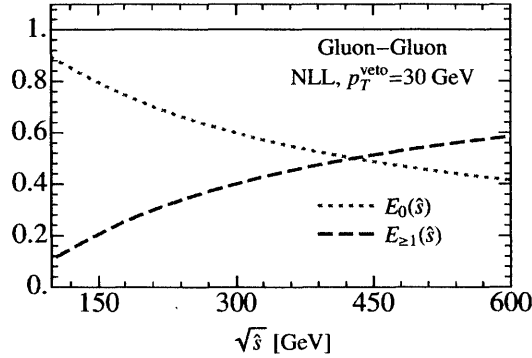
Figure 6-2: The ratios $E_0(\hat{s})$, $E_{\geq 1}(\hat{s})$ for a gluon-gluon initial state, and $p_T^{veto} = 30$ GeV. There is a large migration from the exclusive 0-jet bin to the inclusive 1-jet bin as a function of $\hat{s}$. This phenomenon is important for understanding the impact of jet binning on off-shell cross sections.

In this equation, $\sigma_{\geq 1}(p_T^{veto})$ is the inclusive 1-jet cross section defined by requiring at least one jet with $p_T \geq p_T^{veto}$, σ_0 is the exclusive zero-jet cross section and σ is the inclusive cross section. This relation allows us to discuss the properties of the inclusive 1-jet bin as a function of $\hat{s}$ using the factorization theorem for the exclusive 0-jet cross section. Of particular interest is the split of the total cross section between the exclusive zero-jet bin and the inclusive 1-jet bin, and the migration between the two bins as a function of $\hat{s}$. This relation also implies a correlation between the theory uncertainties for the resummation for the two jet bins, which is important for experimental analyses using jet binning [451].

In Fig. 6-2 we plot $E_0(\hat{s})$, and

$$E_{\geq 1}(\hat{s}) = \left(\frac{d\sigma_{\geq 1}^{NLL}(p_T^{veto})}{d\sqrt{\hat{s}}} \right) \bigg/ \left(\frac{d\sigma}{d\sqrt{\hat{s}}} \right) , \qquad (6.21)$$

as a function of $\hat{s}$ for a gluon-gluon initial state with $p_T^{veto} = 30$ GeV. The behaviour in this plot is of course evident from Fig. 6-1, but it is interesting to interpret it in this fashion: as an $\hat{s}$ dependent migration between jet bins. Although our calculation is only for the inclusive 1-jet bin, the dominant increase will be in the exclusive 1-jet bin.

This migration between the jet bins as a function of $\hat{s}$ is important for the proper un-

derstanding of the off-shell cross section predictions in the presence of jet vetos. For CMS's recent off-shell $H \to ZZ \to 2l2\nu$ analysis, ignoring the VBF category, the events were categorized into exclusive zero jet, and inclusive one jet bins [446], both of which have high sensitivity, due to the clean experimental signal. For the case of $H \to WW$, exclusive 0, 1, and 2 jet bins are used, although the experimental sensitivity is largest in the 0-jet bin, where the backgrounds are minimized.

The effect of the migration is therefore different in the two cases. For $H \to ZZ$, since the backgrounds are easier to control, the jets that migrate from the exclusive 0-jet bin are captured in the inclusive 1-jet bin. Since both are used in the experiment, there is not a significant loss in analysis power. Accurate predictions for the two jet bins should still be used, and the correlations in the theory uncertainties due to resummation should still be treated properly. For the case of $H \to WW$, where the jet veto plays a more essential role in removing backgrounds, the migration causes a loss in sensitivity. For example, the analysis of [434] used the exclusive zero jet bin of $H \to WW$ to bound the Higgs width without a treatment of the $\hat{s}$ dependence induced by the jet veto. This will be discussed further in Sec. 6.5. Calculations for the exclusive 1-jet and 2-jet bins are more difficult. Although NLL resummed results exist for the case $p_T^{jet} \sim \sqrt{\hat{s}}$ [209, 208], the treatment of $p_T^{jet} \ll \sqrt{\hat{s}}$ is more involved [456]. The latter is the kinematic configuration of interest for far off-shell production.

6.3.3 Variation with E_{cm}

Here we comment briefly on the dependence of the exclusive zero jet cross section on the hadronic centre of mass energy, E_{cm}. This is of course of interest as the LHC will resume at $E_{cm} = 13$ TeV in the near future, and Higgs coupling and width measurements are important benchmarks for future colliders at higher energies. Here we only discuss the $\hat{s}$ dependence of the suppression due to the jet veto, the ratio of Eq. (6.19), on E_{cm}. Of course, with an increased E_{cm}, one can more easily achieve higher $\hat{s}$, allowing for off-shell production over a larger range, magnifying the importance of off-shell effects. We will discuss this for the

335
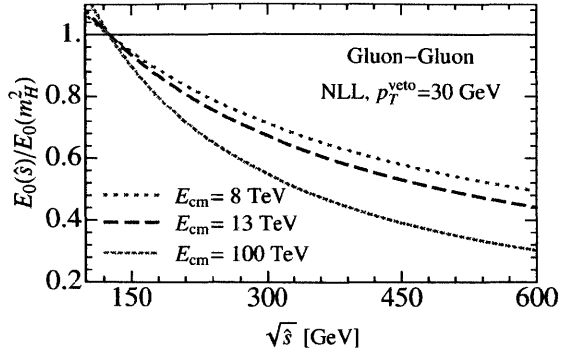
Figure 6-3: A comparison of the effect of the jet veto at $E_{\rm cm} = 8, 13, 100$ TeV for a gluon-gluon initial state, and $p_T^{veto} = 30$ GeV. At higher $E_{\rm cm}$ a larger suppression in the exclusive zero jet bin is observed, due to the larger range of Bjorken x probed.

specific case of $gg \to H \to WW$ in Sec. 6.4.

In Fig. 6-3 we compare the ratio $E_0(\hat{s})/E_0(m_H^2)$ for $E_{\rm cm} = 8, 13, 100$ TeV. As the value of $E_{\rm cm}$ is raised, the $\hat{s}$ dependence of the jet veto suppression systematically increases. Although the effect is relatively small between 8 TeV and 13 TeV, it is significant at 100 TeV. A similar effect was discussed in [457] where the exclusive zero jet fraction for on-shell Higgs production was observed to decrease with increasing $E_{\rm cm}$. Since the Sudakov factor is independent of $E_{\rm cm}$, this difference arises due to the fact that as the $E_{\rm cm}$ is increased, the PDFs are probed over a larger range of Bjorken x, including smaller $x_{a,b}$ values. In the NLL factorization theorem of Eq. (6.17) the PDFs are evaluated at the scale p_T^{veto} instead of at the scale $\hat{s}$. The impact of this change of scales in the PDFs depends on the x values probed, and causes an increasing suppression as $E_{\rm cm}$ is increased.

For the majority of this chapter we will restrict ourselves to $E_{\rm cm} = 8$TeV, although in Sec. 6.4.6 we will further discuss the effect of an increased $E_{\rm cm}$.

6.4 $gg \to H \to WW$: A Case Study

In this section we use $gg \to H \to WW$ to discuss the effect of an exclusive jet veto in more detail. $H \to WW$ is a particularly interesting example to demonstrate the $\sqrt{\hat{s}}$ dependence of

the jet veto suppression since it has a sizeable contribution from far off-shell production [443, 444], and furthermore has interference with continuum $gg \to WW \to l\nu l\nu$ production, which contributes over a large range of $\sqrt{\hat{s}}$ [439, 440, 442]. A jet veto is also required experimentally for this channel due to large backgrounds. For the signal-background interference, we will consider two different Higgs masses, $m_H = 126$ GeV and $m_H = 600$ GeV, which have interference which depend differently on $\sqrt{\hat{s}}$ and therefore cover two interesting scenarios for the different effects that the jet veto can have.

In Sec. 6.4.1 we discuss in detail the hard coefficients, and the matching to SCET. Default parameters are given in Sec. 6.4.2. In Sec. 6.4.3 we use $gg \to H \to WW \to l\nu l\nu$, which can be calculated to NLL and NNLL, to study the convergence in the off-shell region. The extension to NNLL allows us to study the effect of the finite radius of the jet veto. In Sec. 6.4.4 we show results for the NLL resummation for the signal-background interference. Although we are unable to go to NNLL without the NLO hard function for the interference, the results of Sec. 6.4.3 give us confidence that the NLL result is capturing the dominant effects imposed by the jet veto restriction. In Sec. 6.4.5 we consider jet veto suppression in the exclusive zero jet bin as a function of the experimental observable M_T.

6.4.1 Hard Function and Matching to SCET

In this section we discuss the hard function appearing in the SCET factorization theorem, which carries the dependence on the hard underlying process. This is discussed in some detail, as we will be considering signal-background interference, which has not previously been discussed in the language of SCET.

It was shown in [439] that only two Feynman diagram topologies contribute to the process $gg \to \nu_e e^+ \mu^- \bar{\nu}_\mu$ at LO, due to a cancellation between diagrams with an s-channel Z boson. The two diagrams that contribute are the gluon-fusion Higgs diagram, and a quark box diagram for the continuum production, both of which are shown in Fig. 6-4. The $gg \to \nu_e e^+ \mu^- \bar{\nu}_\mu$ cross section consists of Higgs production, the continuum production, and the interference between the two diagrams. Although the interference contribution is small
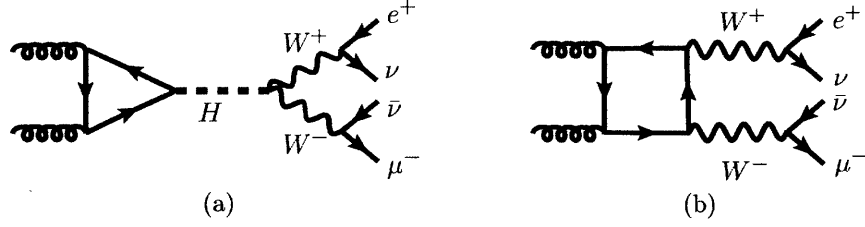
Figure 6-4: LO Higgs mediated, (a), and continuum, (b), diagrams contributing to the process $gg \to l\nu l\nu$. These are matched onto the helicity basis of SCET operators given in Eq. (6.27).

when considering on-shell Higgs production, it becomes important in the off-shell region.

In the effective field theory formalism, these two diagrams are matched onto effective operators in SCET. It is convenient both for understanding the interference, and for comparing with fixed order QCD calculations to work in a helicity and color operator basis in SCET [298, 43]. For this process the color structure is unique, as we are considering the production of a color singlet state from two gluons. We therefore focus on the helicity structure. The helicity of the outgoing leptons is fixed by the structure of the weak interactions, so we need only construct a helicity basis for the incoming gluons. We write the amplitudes for the above diagrams as

$$\mathcal{A}_H(1_g^{h_1}, 2_g^{h_2}, 3_{\nu_e}^-, 4_{\bar{e}}^+, 5_\mu^-, 6_{\bar{\nu}_\mu}^+), \quad \mathcal{A}_C(1_g^{h_1}, 2_g^{h_2}, 3_{\nu_e}^-, 4_{\bar{e}}^+, 5_\mu^-, 6_{\bar{\nu}_\mu}^+) \tag{6.22}$$

where the subscripts H, C denote the Higgs mediated, and continuum box mediated diagrams respectively, and the superscripts denote helicity. In the following we will mostly suppress the lepton arguments, as their helicities are fixed, and focus on the gluon helicities.

Since the SM Higgs boson is a scalar, we have

$$\mathcal{A}_H(1_g^-, 2_g^+) = \mathcal{A}_H(1_g^+, 2_g^-) = 0. \tag{6.23}$$

In this chapter, our focus is on the Higgs production and the signal-background interference. Since there is no interference between distinct helicity configurations, we can therefore also

338

ignore the continuum production diagrams with the $+-$, $-+$ helicity configuration. These do contribute to the background, however their contribution is small compared to the $q\bar{q} \to l\nu l\nu$ process.

The above amplitudes are matched onto operators in the effective theory. The SCET operators at leading power are constructed from collinear gauge-invariant gluon fields [22, 23]

$$\mathcal{B}^{\mu}_{n,\omega\perp} = \frac{1}{g}\left[\delta(\omega + \bar{\mathcal{P}}_n)W^{\dagger}_n(x)i\mathcal{D}^{\mu}_{n\perp}W_n(x)\right] \tag{6.24}$$

where n, $\bar{n}$ are lightlike vectors along the beamline. The collinear covariant derivative is defined as

$$i\mathcal{D}^{\mu}_{n\perp} = \mathcal{P}^{\mu}_{n\perp} + gA^{\mu}_{n\perp}, \tag{6.25}$$

with $\mathcal{P}$ a label operator which extracts the label component of the momentum in the effective theory, and W_n is a Wilson line defined by

$$W_n(x) = \left[\sum_{\text{perms}} \exp\left(-\frac{g}{\mathcal{P}_n}\bar{n} \cdot A_n(x)\right)\right]. \tag{6.26}$$

A helicity basis of SCET operators for the process of interest is given by

$$\mathcal{O}^{++} = \frac{1}{2}\mathcal{B}^a_{n+}\mathcal{B}^a_{\bar{n}+}J_{34-}J_{56-} \tag{6.27}$$

$$\mathcal{O}^{--} = \frac{1}{2}\mathcal{B}^a_{n-}\mathcal{B}^a_{\bar{n}-}J_{34-}J_{56-}, \tag{6.28}$$

where the $1/2$ is a bosonic symmetry factor to simplify matching to the effective theory. We have defined collinear gluon fields of definite helicity by

$$\mathcal{B}^a_{i\pm} = -\epsilon_{\mp\mu}\mathcal{B}^{a,\mu}_{n_i,\omega_i\perp_i}, \tag{6.29}$$

where $\epsilon_{\mp\mu}$ are polarization vectors, as well as leptonic currents of definite helicity

$$J_{ij-} = \epsilon^{\mu}_+(p_i, p_j)\frac{\bar{\psi}_{i-}\psi_{j-}}{\sqrt{2}[ji]}. \tag{6.30}$$

339

In this expression, and in the expressions for the Wilson coefficients given below, we will use the standard spinor helicity notation, with $\langle ij \rangle = \bar{u}_-(p_i)u_+(p_j)$, and $[ij] = \bar{u}_+(p_i)u_-(p_j)$.

Note that we use Hard functions that are fully differential in the leptonic momenta. This allows for realistic experimental cuts on the leptonic phase space to be straightforwardly incorporated.

It is important to note that operators with distinct external helicities do not mix under the SCET RGE at leading power. The jets from the incoming partons, which are described by the beam functions, can only exchange soft gluons, described by the soft function. At leading power, the soft gluons cannot exchange spin, only color, and therefore the RGE can only mix Wilson coefficients in color space, which in this case is trivial. This allows one to consistently neglect the operators $\mathcal{O}^{+-}$, $\mathcal{O}^{-+}$, which would arise from matching the $\mathcal{A}_C(1_g^-, 2_g^+)$, and $\mathcal{A}_C(1_g^+, 2_g^-)$ onto SCET. They do not contribute to the process of interest, and do not mix under the RGE with the operators that do contribute.

We are interested in considering both the direct Higgs production and signal-background interference separately, so it is convenient to maintain this distinction in SCET. Although the SCET operators are the same in both cases, we can separate the Wilson coefficient into a component from the Higgs mediated diagram, and a component from the box mediated continuum diagram. We then have four Wilson coefficients

$$C_{++}^H, \ C_{--}^H, \ C_{++}^C, \ C_{--}^C. \tag{6.31}$$

Since the operators are in a helicity basis, these four Wilson coefficients are simply the finite part of the helicity amplitudes for the given processes (or more specifically for $\overline{\text{MS}}$ Wilson coefficients in SCET are the finite part of the helicity amplitudes computed in pure dimensional regularization). These were computed in [439], and can be obtained from the MCFM code [458]. The Wilson coefficients for the Higgs mediated process depend on the Higgs and W boson widths and masses, as well as the invariants s_{12}, s_{34}, s_{56}. The explicit

340

leading order Wilson coefficients for the Higgs mediated process are given by

$$C_{--}^{H}(m_H, \Gamma_H, s_{12}, s_{34}, s_{56}) = \left(\frac{g_w^4 g_s^2}{16\pi^2}\right) \mathcal{P}_H(s_{12})\mathcal{P}_W(s_{34})\mathcal{P}_W(s_{56})\frac{\langle 12\rangle\langle 35\rangle[64]}{[21]s_{34}s_{56}}F_H(s_{12}), \quad (6.32)$$

$$C_{++}^{H}(m_H, \Gamma_H, s_{12}, s_{34}, s_{56}) = \left(\frac{g_w^4 g_s^2}{16\pi^2}\right) \mathcal{P}_H(s_{12})\mathcal{P}_W(s_{34})\mathcal{P}_W(s_{56})\frac{[12]\langle 35\rangle[64]}{\langle 21\rangle s_{34}s_{56}}F_H(s_{12}), \quad (6.33)$$

where the function $\mathcal{P}_i$ is the ratio of the propagator for the particle species i to that of the photon,

$$\mathcal{P}_i(s) = \frac{s}{s - m_i^2 + i\Gamma_i m_i}. \quad (6.34)$$

We have also used $F_H(s_{12})$ for the usual loop function for gluon-fusion Higgs production

$$F_H(s_{12}) = \sum_{q=t,b} \frac{m_q^2}{s_{12}}\left[2 + \left(\frac{4m_q^2}{s_{12}} - 1\right)g\left(\frac{m_q^2}{s_{12}}\right)\right], \quad (6.35)$$

with

$$g(x) = \begin{cases} \frac{1}{2}\left[\log\left(\frac{1+\sqrt{1-4x}}{1-\sqrt{1-4x}}\right) - i\pi\right]^2 & x < \frac{1}{4} \\ -2\left(\sin^{-1}\left(\frac{1}{2\sqrt{x}}\right)\right)^2 & x \geq \frac{1}{4}. \end{cases} \quad (6.36)$$

The Wilson coefficients C_{++}^{C} and C_{--}^{C} for the box diagram depend on the W mass and width, as well as the kinematic invariants formed by the external momenta. In the presence of massive quarks in the loops, they are extremely lengthy, so we do not reproduce them here. We refer interested readers to [439], and the MCFM code from which we have extracted the required results for our analysis. We have verified that our extracted expressions reproduce quoted numerical results and distributions in [439].

The Hard coefficient, H, appearing in the factorization theorem, Eq. (6.7) is given by the square of the Wilson coefficients:

$$H = |C_{++}^{H} + C_{++}^{C}|^2 + |C_{--}^{H} + C_{--}^{C}|^2 + |C_{+-}^{C}|^2 + |C_{-+}^{C}|^2 \quad (6.37)$$

As is typically done in the case of squared matrix elements, we can separate the hard function into the sum of a hard function for the Higgs mediated process H^H, a hard function for the

341

interference H^{int}, and a hard function for the background arising as the square of Wilson coefficient for the continuum process H^C (which we will not use here). For the first two we have

$$H^H = |C^H_{++}|^2 + |C^H_{--}|^2 \tag{6.38}$$

$$H^{int} = 2\mathrm{Re}\left[C^H_{++}(C^C_{++})^\dagger\right] + 2\mathrm{Re}\left[C^H_{--}(C^C_{--})^\dagger\right]. \tag{6.39}$$

This decomposition allows us to discuss the resummation of the interference and the Higgs mediated processes separately in the effective theory, in a language that is identical to that used in Feynman diagram calculations. In Secs. 6.4.3 and 6.4.4 we will discuss the effect of resummation on both the Higgs mediated contribution and the signal-background interference.

6.4.2 Parameters for Numerical Calculations

For the numerical results, we use the default set of electroweak parameters from MCFM, following [434, 439]:

$$m_W = 80.398\,\mathrm{GeV}, \qquad\qquad m_Z = 91.1876\,\mathrm{GeV},$$

$$\Gamma_W = 2.1054\,\mathrm{GeV}, \qquad\qquad \Gamma_Z = 2.4952\,\mathrm{GeV},$$

$$m_t = 172.5\,\mathrm{GeV}, \qquad\qquad m_b = 4.4\,\mathrm{GeV},$$

$$G_F = 1.16639 \times 10^{-5}\,\mathrm{GeV}^{-2}, \qquad\qquad \sin^2\theta_W = 0.222646,$$

$$\alpha_{e.m.}(m_Z) = \frac{1}{132.338}.$$

We use the following two Higgs mass/width combinations to demonstrate the dependence on the Higgs mass:

$$m_H = 126\,\mathrm{GeV}, \qquad \Gamma_H = 0.004307\,\mathrm{GeV},$$

$$m_H = 600\,\mathrm{GeV}, \qquad \Gamma_H = 122.5\,\mathrm{GeV},$$

where the widths are determined from HDECAY [459]. We use the NLO PDF fit of Martin, Stirling, Thorne and Watt [455] with $\alpha_s(m_Z) = 0.12018$.

The results in this section were obtained using the analytic results for the partonic process documented in [439]. Scalar loop integrals were evaluated using the LoopTools package [460], and phase space integrals were done using the Cuba integration package [461]. For all the results presented in this section, we have integrated over the leptonic phase space, and allow for off-shell vector bosons.

6.4.3   Off-Shell Higgs Production

We begin by studying the effect of the jet veto on far off-shell Higgs production in $gg \rightarrow H \rightarrow WW \rightarrow e^+\nu_e\mu\bar{\nu}_\mu$. While a full analysis of the off-shell region also requires the inclusion of signal-background interference, for which the hard function to NLO is not known, we use the off-shell Higgs mediated process to study the convergence of the resummed predictions. In particular, one is first sensitive to the jet radius at NNLL. The ability to perform the resummation to NNLL for the Higgs mediated signal enables us to assess the convergence of the resummed predictions in the off-shell region. It also allows us to check that the NLL result, which will be used when signal-background interference is included, accurately captures the effect of the jet veto reasonably well. In particular, we will focus on the shape of the differential distribution in $\hat{s}$. As will be discussed in more detail in Sec. 6.5, the procedure for extracting a bound on the Higgs width from the off-shell cross section uses a rescaling procedure to the on-shell cross section. Because of this, the shape, but not the normalization of the distribution is important for an accurate application of this method. Therefore, as in Sec. 6.3 we will rescale the differential cross sections by $E_0(m_H^2)$, allowing us to focus just on the shape.

The NNLL calculation requires the NNLL beam and soft functions, which are known in the literature for a jet veto defined by a cut on p_T [13], as well as the virtual part of the NLO gluon-fusion hard function. The NLO virtual contributions for gluon fusion Higgs production are known analytically with full dependence on the top and bottom quark mass [349], which

343
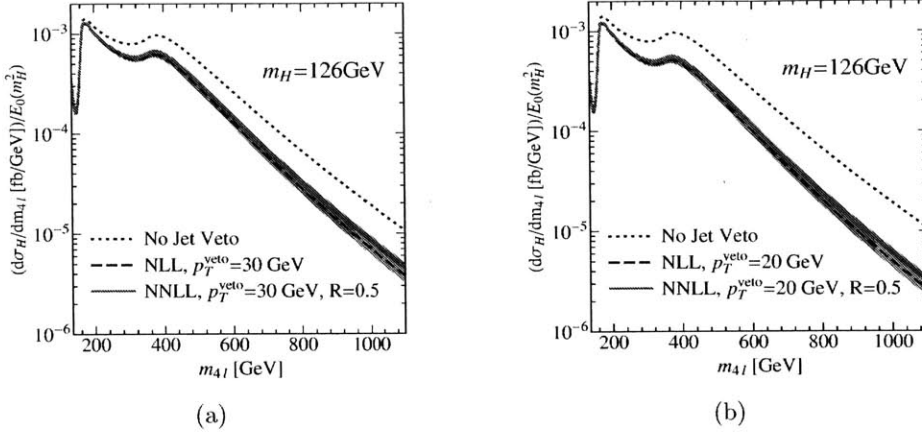
Figure 6-5: The off-shell Higgs cross section in the exclusive zero jet bin for $p_T^{veto} = 30$ GeV in (a), and $p_T^{veto} = 20$ GeV in (b), with $R = 0.5$ in both cases. Results are normalized by the jet veto suppression at the Higgs mass, such that the on-shell cross section is the same in all cases, allowing one to focus on the modification to the shape of the distribution. NLL and NNLL results are similar, with a small modification due to the finite jet radius, which is not present in the NLL calculation.

is necessary, as in the off-shell region one transitions through the $\sqrt{\hat{s}} = 2m_t$ threshold.[2] The NLO hard function is determined by matching onto the gluon-fusion operators in SCET, as discussed in Sec. 6.4.1. We do not include in our calculation the non-singular pieces, as we focus on the region $p_T^{veto} \ll \sqrt{\hat{s}}$, where the singular contributions dominate, and are not interested in the transition to the region $p_T^{veto} \sim \sqrt{\hat{s}}$.

In Fig. 6-5 we plot the resummed distribution, normalized to the jet veto suppression at the Higgs mass: $(d\sigma_0/dm_{4l})/E_0(m_H^2)$, for off-shell $gg \to H \to WW \to e^+ \nu_e \mu \bar{\nu}_\mu$. Note that in the case without a jet veto, the jet veto suppression at the Higgs mass is defined to be 1. We have integrated over the leptonic phase space. Here $m_{4l} = \sqrt{\hat{s}}$ is the invariant mass of the 4 lepton final state. In Fig. 6-5a we use $p_T^{veto} = 30$ GeV, and in Fig. 6-5b we use $p_T^{veto} = 20$ GeV. In both cases, we use a jet radius of $R = 0.5$, as is currently used by the CMS collaboration. The uncertainty bands are rough uncertainty estimates from scale variations by a factor of 2. Note that in the calculation, we use a five flavour scheme, even

---

[2]The analytic NLO virtual corrections were also used in [462] to study the dependence of the jet veto on the b-quark mass for the case of on-shell Higgs production.

above m_t since the difference with using a six flavour coupling is well within our error band.

Figs. 6-5a and 6-5b show a small modification to the differential distribution between NLL and NNLL. This arises primarily due to the clustering logarithms, which introduce dependence on the jet radius, which is not present at NLL. The R dependence reproduces the expected physical dependence of the cross section on R: for a fixed p_T^{veto} cut, the restriction on radiation from the initial partons becomes weaker as the jet radius is decreased, causing a smaller suppression of the cross section. Despite this, the shape is well described by the NLL result. In particular, the NLL result captures the dominant effect of the exclusive jet veto on the off-shell cross section. This is important for the resummation of the interference, considered in Sec. 6.4.4. In this case, higher order results are not available (for some approximate results, see [463]), and therefore one is restricted to an NLL resummation. However, the results of this section demonstrate that the NLL result accurately captures the effects of the jet veto on the shape of the distribution as a function of $\hat{s}$.

6.4.4   Signal-Background Interference

Signal-background interference for the process $gg \to l\nu l\nu$ has been well studied in the literature[439, 440, 442]. The interference comes almost exclusively from the $\sqrt{\hat{s}} > 2M_W$ region. For a light Higgs, $m_H < 2M_W$, this means that the interference comes entirely from $\sqrt{\hat{s}} > m_H$. For a heavy Higgs, the Higgs width is sufficiently large that there are contributions to the interference from a wide range of $\sqrt{\hat{s}}$. The signal-background interference is therefore, in both cases, an interesting process on which to demonstrate the effect of the jet veto.

The NLO virtual corrections are not available for the interference process, restricting the resummation accuracy to NLL. However, as argued in Sec. 6.4.3 if one is interested in the shape of the distribution, and not the normalization, the NLL captures the effects of the jet veto. One thing that cannot be known without a full calculation of the NLO virtual contributions to the interference is if the NLO virtual contributions for the interference are different than for the signal. For the case of interference in $H \to \gamma\gamma$ where they are known,

345

the virtual contributions for the interference were found to be smaller than for the signal [432]. Due to the similar structure of the diagrams for $H \to WW$, the same could certainly be true. However, we expect this to be a minor correction compared to the effects of the jet veto. In particular, we do not expect the K-factor to have strong $\hat{s}$ dependence, which is the important effect captured by the resummation. In this section we use the LO result for $gg \to e\nu\mu\nu$, fully differential in the lepton momenta, which is available in the MCFM code, and is documented in [439].

We begin by reviewing the notation for the signal-background interference in $gg \to e\nu\mu\nu$ at LO following [439]. It is convenient to pull out the dependence on m_H and Γ_H coming from the s-channel Higgs propagator. Defining $\widetilde{C}^H = (\hat{s} - m_H^2 + im_H\Gamma_H)C^H$, we can separate the Hard function for the signal-background interference into its so called "Imaginary" and "Real" contributions:

$$H^{int} = \frac{2(\hat{s} - m_H^2)}{(\hat{s} - m_H^2)^2 + m_H^2\Gamma_H^2}\mathrm{Re}\left[\widetilde{C}^H(C^C)^\dagger\right] + \frac{2m_H\Gamma_H}{(\hat{s} - m_H^2)^2 + m_H^2\Gamma_H^2}\mathrm{Im}\left[\widetilde{C}^H(C^C)^\dagger\right]. \quad (6.40)$$

In Eq. (6.40) there is a sum over helicities of the Wilson coefficients, which for notational convenience has not been made explicit. Note that the imaginary part of the interference is multiplied by an explicit factor of Γ_H, and is therefore negligible for a light Higgs.

The interference without a jet veto is shown in Fig. 6-6a for a 126 GeV Higgs and Fig. 6-6b for a 600 GeV Higgs, as a function of m_{4l}. We have integrated over the phase space of the leptons, including allowing for off-shell vector bosons. The interference is negligible below the $\sqrt{\hat{s}} = 2m_W$ threshold. For the case of $m_H = 126$ GeV the only non-negligible contribution is the real part of the interference above the Higgs pole, which gives a negative contribution to the total cross section. In the case of $m_H = 600$ GeV, there is significant interference both above and below the Higgs pole, and from both the real and imaginary parts. The interference below the pole dominates, leading to a net positive contribution to the total cross section. We have chosen these two Higgs masses, where the interference has a different $\sqrt{\hat{s}}$ dependence, so as to demonstrate the different effects that a jet veto can have on signal-background interference.

346
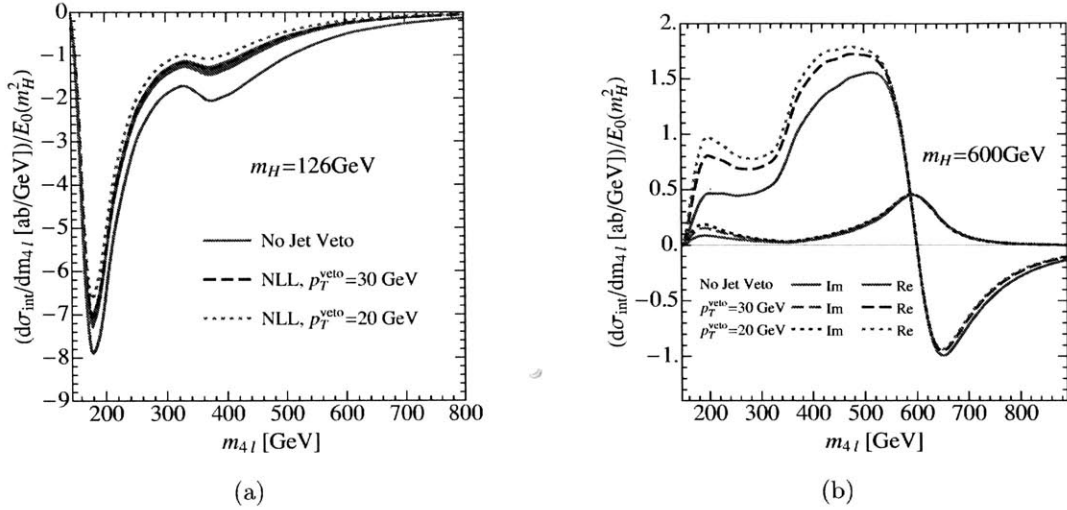
Figure 6-6: Signal-background interference in $gg \to e^+\nu_e\mu\bar{\nu}_\mu$ for (a) $m_H = 126$ GeV, and (b) $m_H = 600$ GeV. NLL predictions are shown for $p_T^{veto} = 20, 30$ GeV, and have been rescaled by the jet veto efficiency at m_H. The size of the signal-background interference relative to the on-shell cross section is enhanced by the jet veto for a heavy Higgs, whereas it is suppressed for a light Higgs. For $m_H = 126$ GeV the jet veto causes a significant reduction of the cross section in the far off-shell region relative to the on-shell cross section.

Fig. 6-6 also shows as a function of $\sqrt{\hat{s}}$ the result for interference including a jet veto of $p_T^{veto} = 20, 30$ GeV with NLL resummation, which can be compared with the interference without a jet veto. To make the interpretation of Fig. 6-6 as simple as possible, we have rescaled the interference by $E_0(m_H^2)$, the jet veto efficiency at m_H. Therefore, enhancements and suppressions in the jet vetoed interference correspond to enhancements and suppressions of the interference relative to the on-shell Higgs contribution when a jet veto is applied. As expected from the discussion in Sec. 6.3, we find a significant suppression of the interference at higher $\sqrt{\hat{s}}$, and this suppression increases with $\sqrt{\hat{s}}$. For $m_H = 126$ GeV, shown in Fig. 6-6a, the interference comes entirely from above $\sqrt{\hat{s}} = m_H$, and is therefore more highly suppressed by the jet veto relative to the on-shell Higgs cross section. However, the situation is quite different for the case of $m_H = 600$ GeV, shown in Fig. 6-6b. Here the dominant contribution to the interference is from the real part in Eq. (6.40), which changes sign at $\sqrt{\hat{s}} = m_H$. The real part of the interference coming from below $\sqrt{\hat{s}} = m_H$ is positive and is partly cancelled

by negative interference from above $\sqrt{\hat{s}} = m_H$ if we integrate over $\hat{s}$. The jet veto suppresses the on-shell cross section and the negative interference from above $\sqrt{\hat{s}} = m_H$ more than the contribution from the positive interference below $\sqrt{\hat{s}} = m_H$, and therefore the jet veto acts to enhance the interference contribution relative to the signal. This enhancement is significant in the case of $m_H = 600$ GeV, as the interference has contributions starting at $m_{4l} \simeq 2m_W$, where the suppression due to the jet veto is smaller. To quantify this further we can consider the effect of the jet veto on the ratio

$$R_I = \frac{\sigma_{H+I}}{\sigma_H}, \qquad (6.41)$$

where σ_{H+I} is the cross section including the signal-background interference, and σ_H is the Higgs mediated cross section. The behaviour of this ratio is different for the two Higgs masses considered. Numerical values of R_I are shown in Table 6.1. The effect of interference for $m_H = 126$ GeV with or without the jet veto is fairly small, and would be made even smaller when cuts are made to eliminate interference. However, for $m_H = 126$ GeV the effect of the jet-veto can also be significantly amplified when cuts are used to maximize sensitivity to the Higgs width. For example, the analysis of [434] considered the region $M_T > 300$ GeV to bound the Higgs width. Since $m_{4l} \geq M_T$, we see from Fig. 6-6a that in this region the effect of the exclusive jet veto is by no means a small effect, giving a suppression of $\sim 1.5 - 2$. A representative error band from scale variation is also shown in Fig. 6-6a. The effect on the derived bound will be discussed in Sec. 6.5.

These two examples demonstrate that a jet veto can have an interesting interplay with signal-background interference, enhancing or suppressing its contribution relative to the Higgs mediated cross section, depending on the particular form of the interference. A detailed understanding of the interference is of phenomenological interest for both a light and heavy Higgs. In the case of $m_H = 126$ GeV, the interference can be efficiently removed by cuts when studying the on-shell cross section [439], but is important for the understanding of the off-shell cross section. In the case of a heavy Higgs, the interference is important for heavy Higgs searches [439, 464, 465], where it is a large effect, and cannot be easily removed by

348

	No Veto	$p_T^{veto} = 30$ GeV	$p_T^{veto} = 20$ GeV
$m_H = 126$ GeV	0.92	0.94	0.95
$m_H = 600$ GeV	1.38	1.49	1.54

Table 6.1: Values of $R_I = \frac{\sigma_{H+I}}{\sigma_H}$ for $m_H = 126$ GeV and $m_H = 600$ GeV for two different values of p_T^{veto}. As is clear from Fig. 6-6, the jet veto causes a suppression of the importance of the interference relative to the Higgs mediated process for a light Higgs, and an enhancement for a heavy Higgs.
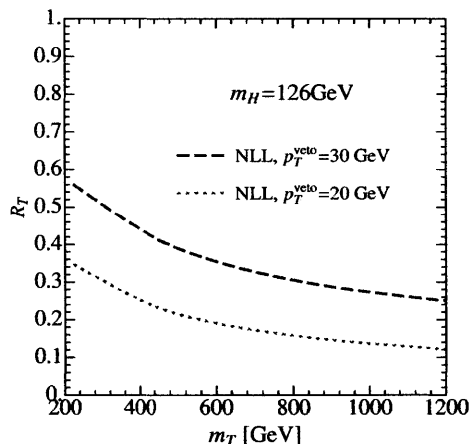
Figure 6-7: Suppression of the exclusive zero jet cross section for off-shell Higgs production as a function of M_T. The Sudakov factor is controlled by $\sqrt{\hat{s}}$, but since $\sqrt{\hat{s}} \geq M_T$, a larger suppression is observed as a function of M_T.

cuts. The effect of the jet veto must therefore be incorporated in such searches.

6.4.5 Suppression as a Function of M_T

We have so far discussed the effect of the jet veto on the cross section as a function of $\sqrt{\hat{s}}$, as the Sudakov factor is explicitly a function of $\sqrt{\hat{s}}$. However, in the case of $H \to WW \to l\nu l\nu$, the total invariant mass of the leptons cannot be reconstructed. A substitute for $\sqrt{\hat{s}}$, used in [439], and which is measured by the CMS and ATLAS collaborations [466, 464, 467] is the transverse mass variable, M_T defined in Eq. (6.5).

Similarly to the ratios considered in Sec. 6.3 of the exclusive zero jet cross section to the

349

total cross section, as a function of $\hat{s}$, in Fig. 6-7, we plot the variable

$$R_T = \left(\frac{d\sigma_0^{NLL}(p_T^{veto})}{dM_T} \right) \bigg/ \left(\frac{d\sigma}{dM_T} \right) , \qquad (6.42)$$

for $gg \to H \to WW \to \bar{\nu}_\mu \mu e^+ \nu_e$ in the far off-shell region. Since M_T is designed as a proxy for $\hat{s}$, the behaviour is as expected from the discussion of Sec. 6.3, however, since $\sqrt{\hat{s}} \geq M_T$, the events contributing at a given M_T all have a larger $\sqrt{\hat{s}}$. Since it is the $\sqrt{\hat{s}}$ that governs the Sudakov suppression due to the jet veto, the suppression due to the jet veto at a given M_T is larger than at the same value of $\sqrt{\hat{s}}$.

We should note that while the values of M_T at which the suppression due to the jet veto becomes significant are larger than is normally considered, or studied experimentally, the authors of [434] show that with an improved understanding of the backgrounds in the ATLAS $N_{jet} = 0$ bin of the WW data, the $M_T > 300$ GeV region, where the jet veto effects are indeed significant, can be used to place a competitive bound on the Higgs width. As will be discussed in Sec. 6.5, their method relies heavily on having an accurate description of the shape of the shape of the M_T distribution, which is modified by the jet veto. This section demonstrates that in the exclusive zero jet bin, there is a suppression by a factor of ~ 2 above $M_T > 300$ GeV, which is a significant effect. This will cause a corresponding weakening of the bound on the Higgs width by a similar factor, which we discuss further in Sec. 6.5.

6.4.6   From 8 TeV to 13 TeV

Since the focus will soon shift to the 13 TeV LHC, in this section we briefly comment on how the effects discussed in the previous sections will be modified at higher E_{cm}. In Sec. 6.3.3 we noted that at higher E_{cm} the jet veto suppression has an increased dependence on $\hat{s}$ due to the larger range of Bjorken x that is probed in the PDFs. The larger range of available x increases the gluon luminosity at high $\hat{s}$ allowing for an increased contribution to the cross section from far off-shell effects [433, 434], and increasing the range over which they

$(\mathrm{d}\sigma_{\mathrm{int}}/\mathrm{d}m_{4l}\ [\mathrm{ab/GeV}])/E_0(m_H^2)$

m_{4l} [GeV]

m_H=126GeV

—— 8 TeV, No Jet Veto

--- 13 TeV, No Jet Veto

······ 13 TeV, p_T^{veto}=30 GeV
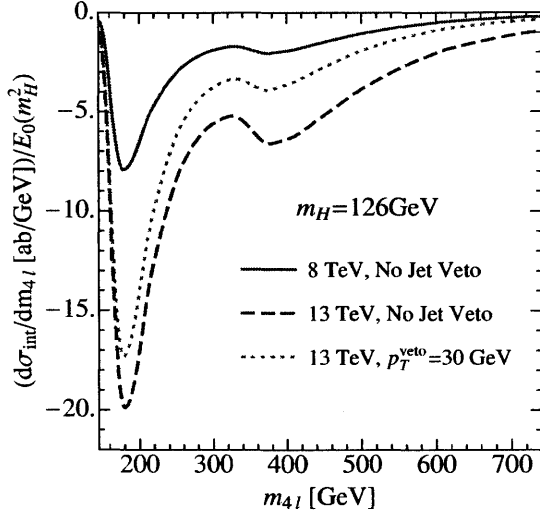
Figure 6-8: A comparison of the signal-background interference, and impact of the jet veto, at 8, 13 TeV. At a higher centre of mass energy there is a larger contribution to the cross section from higher m_{4l}, where the effect of the exclusive jet veto is largest.

contribute, potentially amplifying the effects of the jet veto discussed in the previous sections.

In this section we consider one example to demonstrate this point, the signal-background interference for $m_H = 126$ GeV. The signal-background interference distribution for $E_{\mathrm{cm}} = 8$ and 13 TeV is shown in Fig. 6-8, along with the signal-background interference in the exclusive zero-jet bin at 13 TeV. As was done in Fig. 6-6 for the NLL predictions, we have normalized the distribution by the jet veto suppression at m_H so that the suppression is relative to the on-shell production. The most obvious modification compared with $E_{\mathrm{cm}} = 8$ TeV, is the significant enhancement of the signal-background interference cross section, due to the large enhancement of the gluon luminosity at larger $\hat{s}$. In particular, the contribution to the signal-background interference cross section from the peak at $m_{4l} = 2m_t$ is enhanced at higher E_{cm}, relative to the contribution from $m_{4l} \sim 2m_W$. Since there is a larger relative contribution from higher invariant masses, where the suppression due to the jet veto is larger, the effect of an exclusive jet veto is larger at 13 TeV. This is in addition to the fact that at 13 TeV, the $\hat{s}$ dependence of the suppression due to the jet veto is slightly stronger, as was

351

demonstrated in Sec. 6.3.3.

We again emphasize that when cuts are applied to gain sensitivity to the off-shell region, the effect of the jet veto is not small. In particular, for 13TeV, there is a significant region above $m_{4l} \sim 350$GeV where the suppression due to the jet veto is $\gtrsim 2$, as is seen in Fig. 6-8.

Although we have focused on the effect of an increased centre of mass energy on a particular observable, the conclusions apply generically, for example, for the $H \rightarrow ZZ \rightarrow 4l, 2l2\nu$ channel, which exhibits similar signal-background interference. Indeed as the centre of mass energy is increased one has the ability to probe phenomena over an increasingly large range of $\hat{s}$. This amplifies the effects of off-shell physics, as well as the effect of an exclusive jet veto. These effects will be important in any physics channel for which a jet veto is applied, and for which one is interested in the physics over a range of $\hat{s}$.

6.5   Effect of Jet Vetoes on Higgs Width Bounds

In this section we discuss the impact of jet binning and jet vetoes for the recent program to use the off-shell cross section in $H \rightarrow WW$, ZZ to bound the Higgs width [45, 433, 446, 434]. Although we have focussed on the case of $H \rightarrow WW$, we will review also the strategy for $H \rightarrow ZZ$ which is similar, also exhibiting a large contribution from the far off-shell cross section, as well as signal-background interference analogous to that in $H \rightarrow WW$. We first discuss the procedure used to bound the Higgs width, and then relate it to our discussion of the suppression of the off-shell cross section in the exclusive zero jet bin. Our focus will be on the effect of the jet vetoes, rather than carrying out a complete numerical analysis. In particular the proper incorporation of backgrounds, and additional experimental cuts is beyond the scope of this chapter.

The method used to bound the Higgs width in Refs. [45, 433, 434] can be phrased in a common language for both $H \rightarrow ZZ$, WW. It is based on the different scalings of the on-shell, off-shell and interference contributions to the Higgs cross section, as discussed in

Sec. 6.1. Recalling the scaling from Eq. (6.4), the total cross section can be written as

$$\sigma_{H+I} = A + B \left(\frac{\Gamma_H}{\Gamma_H^{SM}} \right) + C \sqrt{\frac{\Gamma_H}{\Gamma_H^{SM}}}, \qquad (6.43)$$

where the coefficients A, B, C correspond to on-shell, off-shell Higgs mediated, and signal-background interference contributions, respectively. The coefficients depend strongly on the set of cuts that are applied. To extract a bound on the Higgs width, the procedure is as follows. First, one determines a normalization factor between the experimental data and theoretical prediction, which is as independent as possible of the Higgs width. This can be done for WW by using a strict M_T cut, for example $0.75m_H \leq M_T \leq m_H$, and for ZZ by using a strict cut on $\hat{s}$. In both cases, this essentially eliminates the coefficients B, C corresponding to the off-shell production and interference. Once this normalization factor is determined, it is scaled out from the entire differential distribution, so that we now must consider the ratio of offshell and onshell production cross sections. One can then compute the predicted number of events above some M_T, or $\hat{s}$ value, for example, M_T, $\sqrt{\hat{s}} \geq 300$ GeV. In this region, the interference and off-shell production dominate the cross section, so that the coefficients B, C are significant, and the expected number of events is sensitive to the Higgs width, as can be seen from the scalings in Eq. (6.43). By comparing with the number of events observed by the experiment in this region, one can place a bound on the Higgs width.

This method relies on the ability to normalize the theoretical prediction to data in the low M_T, or $\hat{s}$ region, which is insensitive to the dependence on the Higgs width, and then use the same normalization in the high M_T, or $\hat{s}$ region where there is a large sensitivity to the Higgs width through off-shell production and signal-background interference. However, to be able to do this, one needs to have an accurate theory prediction for the shape of the M_T, or $\hat{s}$ distribution, particularly in the high M_T, or $\hat{s}$ region.

As we have seen, the jet veto significantly modifies the shape of the M_T, or $\hat{s}$ distribution, causing it to fall off more rapidly at high M_T, or $\hat{s}$. Often we presented our results by normalizing the offshell cross sections to the cross section at the Higgs mass. Given the

agreement between theory and experiment at m_H, this normalization corresponds exactly to what is done if the theory prediction is normalized to the experimental data in the on-shell region, and therefore shows the extent to which a prediction without the inclusion of a jet veto overestimates the contribution to the cross section at high M_T or $\hat{s}$ compared with that in the exclusive zero jet bin.

In both the $H \to ZZ$, and $H \to WW$ analyses, jet vetoes or jet binning are used, so it is interesting to consider how they will effect the width bounds. Their use in the two channels is quite different so we will discuss them separately.

For the case of $H \to WW$ the jet veto plays an important role because the exclusive zero jet bin dominates the sensitivity, so the jet veto has a large impact on the potential Higgs width bound. This is because effectively it is more difficult to recover the jets which migrate out of the zero jet bin. The plots of the off-shell distributions in Sec. 6.4 show the extent to which a prediction without the inclusion of a jet veto overestimates the contribution to the cross section at high M_T. This will lead to a weakening on the bound of the Higgs mass, compared with a calculation that does not incorporate the effect of the jet veto. For example, in [434], which first proposed the use of the $H \to WW$ channel, the estimated sensitivity was derived by comparing an inclusive calculation for the off-shell cross section with data in the exclusive zero jet bin. Here the effect of the restriction to the zero jet bin is not small, and will worsen the bounds by a factor of ~ 2, as can be seen by the suppression of the far off-shell cross section in the exclusive zero jet bin, shown in Sec. 6.4. In an analogous experimental analysis this Sudakov suppression from the jet veto will be accounted for up to some level of precision by the use of a parton shower. Because this is such a large effect, we believe that an experimental analysis of the high M_T region performed to bound the Higgs width, would benefit from using an analytic calculation of the jet veto suppression in the exclusive zero jet bin, instead of relying on the parton shower. We have demonstrated that the resummation, including the signal-background interference, can be achieved to NLL. Once the NLO virtual corrections are calculated for the interference, these results can also easily be extended to NNLL.

In the $H \to ZZ \to 2l2\nu$ channel the situation is different, as the jet binning procedure is used to optimize sensitivity, splitting the data into exclusive 0-jet and inclusive 1-jet categories with comparable bounds coming from each category [446]. Because the inclusive 1-jet channel is still experimentally clean, the large migration to the inclusive 1-jet bin discussed in Sec. 6.3.2 should have a small (or no) impact on the width bounds derived from the combination of the two channels in $H \to ZZ$. A proper treatment of the migration of events with changing $\hat{s}$ is still important when considering any improvement that is obtained by utilizing jet binning. The analytic results for the Sudakov form factor discussed here for $H \to WW$ could be utilized for jet bins for $H \to ZZ$ in a straightforward manner.

6.6 Conclusions

In this chapter a study of the effect of jet vetoes on off-shell cross sections was made. A factorization theorem in SCET allowed us to analytically treat the summation of Sudakov logarithms, and make a number of general statements about the effect of the jet veto. In particular, the restriction on radiation imposed by the jet veto causes a suppression to the exclusive 0-jet cross section, and correspondingly an enhancement of the inclusive 1-jet cross section, which depends strongly on $\hat{s}$. For gluon initiated processes the $\hat{s}$ dependence of the suppression is greater than for quark initiated processes, which is important for channels where the signal is dominated by one production channel, and the background by another.

The fact that the jet veto suppression is $\hat{s}$ dependent has interesting effects on differential distributions in $\hat{s}$, as well as on signal-background interference. To demonstrate these effects, we considered the $gg \to H \to WW$ process, which has large off-shell contributions as well as signal-background interference. We performed an NLL resummation for the $gg \to H \to WW \to l\nu l\nu$ process, including a discussion of the resummation for the signal-background interference, for $m_H = 126$ GeV, and $m_H = 600$ GeV. These two examples demonstrated that depending on the structure of the interference, the jet veto can either enhance or suppress interference effects relative to the on-shell production. For a low mass Higgs a suppression is observed, while for a high mass Higgs there is a significant enhancement in the interference.

355

These effects must be properly incorporated in high mass Higgs searches that use jet vetoes.

The modification of differential distributions in $\hat{s}$ or M_T due to the $\hat{s}$ dependence of the jet veto suppression is particularly relevant to a recent program to bound the Higgs width using the off-shell cross section [45, 433, 434, 446]. In particular, for the $H \to WW$ channel, where an exclusive 0-jet veto is imposed to mitigate large backgrounds, the jet veto weakens the bound on the Higgs width by a factor of ~ 2 compared to estimates without accounting for the jet veto. Furthermore, since the suppression in the exclusive 0-jet bin corresponds to an enhancement in the inclusive 1-jet bin, and the migration is significant as a function of $\sqrt{\hat{s}}$ a proper understanding of the effect of the jet veto is crucial for experimental analyses which use jet binning. This migration may for example play some role in $H \to ZZ \to 2l2\nu$, which was recently used by CMS to place a bound on the off-shell Higgs width, and which uses jet binning in exclusive 0-jet and inclusive 1-jet bins [446].

We presented a factorization theorem in SCET which allows for the resummation of large logarithms of $\sqrt{\hat{s}}/p_T^{veto}$, including for the signal-background interference, in a systematically improveable manner. This allows for the analytic study of the effect of the jet veto on the exclusive 0-jet and inclusive 1-jet bins, including the correlations in their theory uncertainties. A complete NNLL calculation would require the calculation of the NLO virtual corrections to the interference, but would allow for the analytic incorporation of jet radius effects, and would place the study of the off-shell cross section on a firmer theoretical footing. Furthermore, since our hard functions are fully differential in leptonic momenta, realistic experimental cuts on the leptonic phase space can be easily implemented. We leave a more detailed investigation, including the treatment of such cuts, and a calculation of the effect of the jet veto on the backgrounds, for future study.

With the LHC beginning its 13 TeV run in the near future, the importance of the effects discussed in this chapter will be amplified. As the centre of mass energy is raised, the range of $\hat{s}$ which can be probed increases. This typically increases the importance of off-shell effects, as well as the impact of the jet veto, which is essential for an accurate description of the differential distribution in $\hat{s}$. In general a proper theoretical understanding of jet vetoes

and jet binning for large $\hat{s}$ can be achieved through resummation, and is important when theoretical cross sections are needed for the interpretation of experimental results.

Chapter 7

Conclusions

In this thesis I have discussed applications of effective field theories, in particular SCET, to describing QCD at the LHC. In particular, I have focused on the development of extensions of SCET to describe additional hierarchical scales present in jet substructure, as well as formalisms in SCET to facilitate the description of more complicated hard scattering processes.

In the area of jet substructure I have introduced new observables to distinguish hadronically decaying boosted electroweak resonances from QCD background jets. These observables are theoretically clean, and offer advantages over previously proposed observables. Motivated by these observables, I have proposed effective field theory descriptions of subjet configurations allowing for a systematic resummation and factorization of jet substructure observables, and performed the first calculation of such an observable. For the calculation of precision cross sections at the LHC, I have developed a helicity operator formalism in SCET to facilitate the matching of fixed order results onto the effective theory. This greatly simplifies the analysis of final states with multiple particles.

The effective theory tools developed in this thesis should aid in furthering our understanding of fundamental physics using the Large Hadron Collider over the coming decades.

Appendix A

One-Loop Calculations for D_2

A.1   Conventions and SCET Notation

In the body of the text we have presented the required factorization theorems for studying the two-prong substructure of jets using the D_2 observable. Although all the factorization theorems were presented, only heuristic descriptions of the functions appearing in the factorization theorems were presented in an attempt to appeal to a broader audience, and so as to not distract the reader with technical complications. In these appendices, we give the operator definitions of the functions appearing in the factorization theorems of Sec. 4.3, and calculate the functions to one-loop accuracy.

In this appendix we begin by summarizing some notation and conventions. The factorization theorems presented in this chapter are formulated in the language of SCET [22, 23, 24, 25]. We assume that the reader has some familiarity with the subject, and will only define our particular notation, and review the definition for common SCET objects. We refer readers unfamiliar with SCET to the reviews [468, 469].

SCET is formulated as a multipole expansion in the momentum components along the jet directions. Since we take the jet directions to be lightlike, it is convenient to work in

terms of light-cone coordinates. We define two light-cone vectors

$$n^\mu = (1, \vec{n}), \qquad \mathbf{n}^\mu = (1, -\vec{n}), \qquad (A.1)$$

with $\vec{n}$ a unit three-vector, which satisfy the relations $n^2 = \mathbf{n}^2 = 0$ and $n \cdot \mathbf{n} = 2$. We can then write any four-momentum p as

$$p^\mu = \mathbf{n} \cdot p \, \frac{n^\mu}{2} + n \cdot p \, \frac{\mathbf{n}^\mu}{2} + p_{n\perp}^\mu. \qquad (A.2)$$

A particle in the n-collinear sector has momentum p close to the $\vec{n}$ direction, so that its momentum scales like $(n \cdot p, \mathbf{n} \cdot p, p_{n\perp}) \sim \mathbf{n} \cdot p \; (\lambda^2, 1, \lambda)$, with $\lambda \ll 1$ a small parameter. The parameter λ is a generic substitute for the power counting parameters in the different factorization theorems presented in Sec. 4.3, and since our factorization theorems involve multiple scales, there are generically multiple distinct λs.

In the effective field theory, the momentum of the particles in the n-collinear sector are multipole expanded, and written as

$$p^\mu = \tilde{p}^\mu + k^\mu = \mathbf{n} \cdot \tilde{p} \, \frac{n^\mu}{2} + \tilde{p}_{n\perp}^\mu + k^\mu, \qquad (A.3)$$

where $\mathbf{n} \cdot \tilde{p}$ and $\tilde{p}_{n\perp}$ are large momentum components, which label fields, while k is a small residual momentum, suppressed by powers of λ. This gives rise to an effective theory expansion in powers of λ.

SCET fields for quarks and gluons in the n-collinear sector, $\xi_{n,\tilde{p}}(x)$ and $A_{n,\tilde{p}}(x)$, are labeled by the lightlike vector of their collinear sector, n and their large momentum $\tilde{p}$. We will write the fields in a mixed position space/momentum space notation, using position space for the residual momentum and momentum space for the large momentum components. The residual momentum dependence can be extracted using the derivative operator $i\partial^\mu \sim k$, while the large label momentum is obtained from the momentum label operator $\mathcal{P}_n^\mu$.

Operators and matrix elements in SCET are constructed from collinearly gauge-invariant

362

quark and gluon fields, defined as [22, 23]

$$\chi_{n,\omega}(x) = \left[\delta(\omega - \overline{\mathcal{P}}_n)\, W_n^\dagger(x)\, \xi_n(x)\right] , \tag{A.4}$$

$$\mathcal{B}^\mu_{n,\omega\perp}(x) = \frac{1}{g}\left[\delta(\omega + \overline{\mathcal{P}}_n)\, W_n^\dagger(x)\, iD^\mu_{n\perp} W_n(x)\right] . \tag{A.5}$$

The $\perp$ derivative in the definition of the SCET fields is defined using the label momenta operator as

$$iD^\mu_{n\perp} = \mathcal{P}^\mu_{n\perp} + gA^\mu_{n\perp} , \tag{A.6}$$

and

$$W_n(x) = \left[\sum_{\text{perms}} \exp\left(-\frac{g}{\overline{\mathcal{P}}_n}\, \mathbf{n}\cdot A_n(x)\right)\right] , \tag{A.7}$$

is a Wilson line of n-collinear gluons. We use the common convention that the label operators in the definition of the SCET fields only act inside the square brackets. Although the Wilson line $W_n(x)$ is a non-local operator, it is localized with respect to the residual position x, and we can therefore treat $\chi_{n,\omega}(x)$ and $\mathcal{B}^\mu_{n,\omega}(x)$ as local quark and gluon fields when constructing operators. The operator definitions for jet functions in these appendices are given in terms of these collinear gauge invariant quark and gluon SCET fields.

Our operator definitions will also involve matrix elements of eikonal Wilson lines, which arise from the soft-collinear factorization through the BPS field redefinition at the Lagrangian level [24]. The Wilson lines extend from the origin to infinity along the direction of a lightlike vector, q, specifying their directions. Explicitly

$$S_q = \mathbf{P}\exp\left(ig\int\limits_0^\infty ds\, q\cdot A(x + sq)\right) . \tag{A.8}$$

Here $\mathbf{P}$ denotes path ordering, and A is the appropriate gauge field for any sector which couples eikonally to a collinear sector with label q (for example collinear-soft, soft, boundary soft), and the color representation has been suppressed. All Wilson lines are taken to be outgoing, since we consider the case of jet production from e^+e^- collisions.

Throughout this chapter we have considered the production of two jets, one of which has a possible two-prong substructure, in an e^+e^- collider. This implies the presence of at most three Wilson lines in the soft or collinear soft function. With only three Wilson lines, all possible color structures can be written as a sum of color-singlet traces. In the more general case, with more than three Wilson lines, the soft function is a color matrix which must be traced against the hard functions, which are also matrices in color space, appearing in the factorization theorem for the cross section (see e.g. Refs. [167, 301] for more details).

In App. A.2 through App. A.5 we will give operator definitions for the functions appearing in the factorization theorems in terms of matrix elements of the SCET operators, $\chi_{n,\omega}(x)$ and $\mathcal{B}_{n,\omega}^\mu(x)$, as well as products of soft Wilson lines. These matrix elements can be calculated using the leading power SCET Lagrangian, which can be found in Refs. [22, 23, 24, 25], or by using eikonal Feynman rules in the soft functions, and known results for the splitting functions to calculate the jet functions [330]. We will use the latter approach, as it considerably simplifies the calculations at one-loop.

A.2 One Loop Calculations of Collinear Subjets Functions

In this appendix we collect the calculations relevant to the calculation in the collinear subjets region of phase space, and explicitly show the cancellation of anomalous dimensions. The calculation follows closely that of Ref. [105], with the exception of the form of the measurement function. Nevertheless, the calculation is presented in detail, as the SCET$_+$ effective theory has not been widely used.

Kinematics and Notation

For our general kinematic setup, we will denote by Q the center of mass energy of the e^+e^- collisions, so that $Q/2$ is the energy deposited in a hemisphere. i.e. the four-momenta of the

two hemispheres are

$$p_{\text{hemisphere}_1} = \left(\frac{Q}{2}, \vec{p}_1\right), \qquad p_{\text{hemisphere}_2} = \left(\frac{Q}{2}, -\vec{p}_1\right) \qquad\text{(A.9)}$$

so

$$s = Q^2. \qquad\qquad\qquad\qquad\text{(A.10)}$$

We will also denote the energy in a jet at intermediate stages of the calculation by E_J, but we will write our final results in terms of Q.

We work in the region where one hemispherical jet splits into two hard subjets, assume the power counting $z \sim \frac{1}{2}$, with z being the energy fraction of one of the jets. We further assume the power counting relations between the energy correlation functions valid in the collinear subjets region, as discussed in Sec. 4.2.2. We adopt the following notation to describe the kinematics of the subjets

Subjet a,b momenta:	p_a, p_b	(A.11)
Subjet a,b spatial directions:	$\hat{n}_a, \hat{n}_b$	(A.12)
Thrust axis:	$\hat{n} = \dfrac{\hat{n}_a + \hat{n}_b}{|\hat{n}_a + \hat{n}_b|}$	(A.13)
Light-cone vectors:	$n = (1, \hat{n}),\ \bar{n} = (1, -\hat{n}),$	
	$n_{a,b} = (1, \hat{n}_{a,b}),\ \bar{n}_{a,b} = (1, -\hat{n}_{a,b}).$	(A.14)

In the collinear soft region of phase space, we have $n_a \cdot n_b \ll 1$. When performing expansions, we can work to leading order in $n_a \cdot n_b$, and must use a consistent power counting. It is therefore useful to collect some kinematic relations between vectors which are valid at leading power. These will be useful for later evaluations of the measurement function and integrand at leading power. These kinematics satisfy the following useful relations

$$n \cdot n_a = n \cdot n_b = \frac{n_a \cdot n_b}{4} \qquad\qquad\text{(A.15)}$$

$$\bar{n} \cdot n_a = \bar{n} \cdot n_b = 2 \,, \tag{A.16}$$

$$n_{\perp a,b} \cdot \bar{n}_{\perp a,b} = -n_{\perp a,b} \cdot n_{\perp a,b} = \hat{n}_{\perp a,b} \cdot \hat{n}_{\perp a,b} = \frac{n_a \cdot n_b}{2} \,. \tag{A.17}$$

For a particle with the power counting of collinear sector a or b, we have the following simplified relations

$$p_a \sim \frac{1}{2}(\bar{n} \cdot p_a)n_a, \quad p_b \sim \frac{1}{2}(\bar{n} \cdot p_b)n_b \,, \tag{A.18}$$

$$p_a^0 \sim \frac{1}{2}(\bar{n} \cdot p_a), \quad p_b^0 \sim \frac{1}{2}(\bar{n} \cdot p_b) \,, \tag{A.19}$$

which are true to leading order in the power counting. Finally, we label the energy fractions carried in each subjet by

$$z_{a,b} = \frac{2p_{a,b}^0}{Q} = \frac{\bar{n} \cdot p_{a,b}}{Q} \,, \tag{A.20}$$

where the second relation is true to leading power.

The value of $e_2^{(\alpha)}$ is given to leading power by the subjet splitting

$$e_2^{(\alpha)} = \frac{1}{E_J^2} E_a E_b \left(\frac{2p_a \cdot p_b}{E_a E_b} \right)^{\alpha/2} \tag{A.21}$$

$$= 2^{\alpha/2} z_a z_b \left(n_a \cdot n_b \right)^{\alpha/2} \,. \tag{A.22}$$

In the collinear soft region of phase space, the 3-point energy correlation function is dominated by the correlation between two particles in different subjets, with a third collinear, soft, or collinear-soft particle. Depending on the identity of the third particle, the power counting of the observable is different. We begin by collecting expressions for the $e_3^{(\alpha)}$ observable for a single soft, collinear-soft, or collinear emission, which will be required for the one-loop calculations.

For three emissions, with momenta k_1, k_2, k_3, the general expression for the three point

energy correlation function is

$$e_3^{(\alpha)} = \frac{1}{E_J^3} k_1^0 k_2^0 k_3^0 \left(\frac{2k_1 \cdot k_2}{k_1^0 k_2^0} \right)^{\alpha/2} \left(\frac{2k_1 \cdot k_3}{k_1^0 k_3^0} \right)^{\alpha/2} \left(\frac{2k_2 \cdot k_3}{k_2^0 k_3^0} \right)^{\alpha/2} . \tag{A.23}$$

For an emission collinear with one of the subjets, where we have the splitting $p_{a,b} \to k_1 + k_2$, we can write $e_3^{(\alpha)}$ entirely in terms of $k_1 \cdot k_2$, $n_a \cdot n_b$, and $\bar{n}_a \cdot k_{1,2}$, because there is a hierarchy between the opening angle of the dipole, and the opening angle of the splitting. At leading power it is given by

$$e_3^{(\alpha)} =_{k_1,k_2 \| n_a} 2^{5\alpha/2} z_b (n_a \cdot n_b)^\alpha \left(\frac{k_1 \cdot k_2}{Q^2} \right)^{\frac{\alpha}{2}} \left(\frac{\bar{n}_a \cdot k_1}{Q} \right)^{1-\frac{\alpha}{2}} \left(\frac{\bar{n}_a \cdot k_2}{Q} \right)^{1-\frac{\alpha}{2}} , \tag{A.24}$$

$$e_3^{(\alpha)} =_{k_1,k_2 \| n_b} 2^{5\alpha/2} z_a (n_a \cdot n_b)^\alpha \left(\frac{k_1 \cdot k_2}{Q^2} \right)^{\frac{\alpha}{2}} \left(\frac{\bar{n}_b \cdot k_1}{Q} \right)^{1-\frac{\alpha}{2}} \left(\frac{\bar{n}_b \cdot k_2}{Q} \right)^{1-\frac{\alpha}{2}} . \tag{A.25}$$

For a soft emission off of the dipole, with momentum k, which cannot resolve the opening angle of the dipole, we have

$$n_a \cdot k \to n \cdot k , \qquad n_b \cdot k \to n \cdot k , \tag{A.26}$$

at leading power. We then find

$$e_3^{(\alpha)} = 2^{3\alpha/2+1} z_a z_b (n_a \cdot n_b)^{\alpha/2} \left(\frac{\bar{n} \cdot k + n \cdot k}{2Q} \right)^{1-\alpha} \left(\frac{n \cdot k}{Q} \right)^\alpha , \tag{A.27}$$

where we have used the full expression for the energy of the soft particle, as it is not boosted.

For a third collinear-soft emission k off of the $p_{a,b}$ partons, for which there is no hierarchy between the opening angle of the dipole and the opening angle of the emission (i.e. a collinear soft emission), $e_3^{(\alpha)}$ is given by

$$e_3^{(\alpha)} = 2^{3\alpha/2+1} z_a z_b (n_a \cdot n_b)^{\alpha/2} \left(\frac{\bar{n} \cdot k}{2Q} \right)^{1-\alpha} \left(\frac{n_a \cdot k}{Q} \right)^{\alpha/2} \left(\frac{n_b \cdot k}{Q} \right)^{\alpha/2} . \tag{A.28}$$

For the SCET operators involved in the matching calculation, we follow the notation of

Ref. [105], defining

$$\mathcal{O}_2 = \bar{\chi}_n Y_n^\dagger \Gamma Y_{\bar{n}} \chi_{\bar{n}} \,, \tag{A.29}$$

which is the usual SCET operator for $e^+ e^- \to$ dijets, and

$$\mathcal{O}_3 = \bar{\chi}_{n_a} \mathcal{B}^A_{\perp n_b} \left[X^\dagger_{n_a} X_{n_b} T^A X_{n_b} V_{\bar{n}} \right]_{ij} \left[Y_n^\dagger Y_{\bar{n}} \right]_{jk} \Gamma \left[\chi_{\bar{n}} \right]_k \,, \tag{A.30}$$

which is the SCET$_+$ operator describing the production of the collinear subjets. Throughout this section, we will not be careful with the Dirac structure of the operators, as it is largely irrelevant to our discussion. With this in mind, we have not made the Lorentz indices explicit on the operators. Here we have chosen to write the Wilson line corresponding to the gluon in the fundamental representation. Note that the two stage matching onto SCET$_+$ makes it clear that the partonic configuration in which the two collinear subjets are both quarks is power suppressed. In the operators $\mathcal{O}_2, \mathcal{O}_3$, we have used Y to denote soft Wilson lines, and X, V to denote collinear-soft Wilson lines. In the definitions of the factorized functions below, we will refer to all Wilson lines as S, as after factorization, no confusion can arise.

Definitions of Factorized Functions

The functions appearing in the collinear subjets factorization theorem of Eq. (4.18) have the following SCET operator definitions:

- Hard Matching Coefficient for Dijet Production

$$H\left(Q^2, \mu\right) = \left| C\left(Q^2, \mu\right) \right|^2 \,, \tag{A.31}$$

where $C\left(Q^2, \mu\right)$ is the Wilson coefficient obtained from matching the full theory QCD current $\bar{\psi} \Gamma \psi$ onto the SCET dijet operator $\bar{\chi}_n \Gamma \chi_{\bar{n}}$

$$\langle q\bar{q} | \bar{\psi} \Gamma \psi | 0 \rangle = C\left(Q^2, \mu\right) \langle q\bar{q} | \mathcal{O}_2 | 0 \rangle \,. \tag{A.32}$$

368

When accounting for the Lorentz structure, there is a contraction with the leptonic tensor, which we have dropped for simplicity. See Ref. [167] for a detailed discussion.

- Hard Splitting Function:

$$H_2\left(e_2^{(\alpha)}, z_a, \mu\right) = \left|C_2\left(e_2^{(\alpha)}, z_a, \mu\right)\right|^2 , \tag{A.33}$$

where $C_2\left(e_2^{(\alpha)}, z_a, \mu\right)$ is the Wilson coefficient in the matching from $\mathcal{O}_2$ to $\mathcal{O}_3$, namely the relation between the following matrix elements

$$\langle q\bar{q}g|\mathcal{O}_2|0\rangle = C_2\left(e_2^{(\alpha)}, z_a, \mu\right)\langle q\bar{q}g|\mathcal{O}_3|0\rangle . \tag{A.34}$$

- Jet Function:

$$J_{n_{a,b}}\left(e_3^{(\alpha)}\right) = \tag{A.35}$$

$$\frac{(2\pi)^3}{C_F}\text{tr}\langle 0|\frac{\slashed{\bar{n}}_{a,b}}{2}\chi_{n_{a,b}}(0)\delta(Q - \bar{n}_{a,b}\cdot\mathcal{P})\delta^{(2)}(\vec{\mathcal{P}}_\perp)\delta\left(e_3^{(\alpha)} - \mathbf{E_3}^{(\alpha)}\right)\bar{\chi}_{n_{a,b}}(0)|0\rangle$$

For simplicity, we have given the definition of the quark jet function. The gluon jet function is defined identically but with the SCET collinear invariant gluon field, $\mathcal{B}_{n_{a,b},\perp}$, instead of the collinear invariant quark field.

- Soft Function:

$$S_{n\bar{n}}\left(e_3^{(\alpha)}; R\right) = \frac{1}{C_A}\text{tr}\langle 0|T\{S_n S_{\bar{n}}\}\delta\left(e_3^{(\alpha)} - \Theta_R\mathbf{E_3}^{(\alpha)}\right)\bar{T}\{S_n S_{\bar{n}}\}|0\rangle \tag{A.36}$$

- Collinear-Soft Function:

$$S_{n_a n_b \bar{n}}\left(e_3^{(\alpha)}\right) = \text{tr}\langle 0|T\{S_{n_a}S_{n_b}S_{\bar{n}}\}\delta\left(e_3^{(\alpha)} - \mathbf{E_3}^{(\alpha)}\right)\bar{T}\{S_{n_a}S_{n_b}S_{\bar{n}}\}|0\rangle \tag{A.37}$$

In each of these definitions, we have defined an operator, $\mathbf{E_3}^{(\alpha)}$, which measures the contribution to $e_3^{(\alpha)}$ from final states, and must be appropriately expanded following the power

369

counting of the sector on which it acts, as was shown explicitly in Eq. (A.24), Eq. (A.27), and Eq. (A.28). These operators can be written in terms of the energy-momentum tensor of the full or effective theory [470, 471, 472, 473], but we can simply view them as returning the value of $e_3^{(\alpha)}$ as measured on a particular perturbative state. The soft function is also sensitive to the jet function definition, which is included through the operator Θ_R. To simplify the notation, we have strictly speaking only defined in the in-jet contribution to the soft function. Additionally, we assume that some IRC safe observable is also measured in the out-of-jet region, although this will play little role in our discussion, so we have not made it explicit.

Hard Matching Coefficient for Dijet Production

The hard matching coefficient for dijet production, $H(Q^2, \mu)$, appears in the factorization theorems in each region of phase space. $H(Q^2, \mu)$ is the well known hard function for the production of a $q\bar{q}$ pair in e^+e^- annihilation. It is defined by

$$H\left(Q^2, \mu\right) = \left| C\left(Q^2, \mu\right) \right|^2 , \tag{A.38}$$

where $C\left(Q^2, \mu\right)$ is the Wilson coefficient obtained from matching the full theory QCD current $\bar{\psi}\Gamma\psi$ onto the SCET dijet operator $\bar{\chi}_n\Gamma\chi_{\bar{n}}$. This Wilson coefficient is well known (see, e.g., Refs. [474, 475, 167, 105]), and is given at one-loop by

$$C(Q^2, \mu) = 1 + \frac{\alpha_s(\mu)\, C_F}{4\pi} \left(-\log^2\left[\frac{-Q^2}{\mu^2}\right] + 3\log\left[\frac{-Q^2}{\mu^2}\right] - 8 + \frac{\pi^2}{6} \right) . \tag{A.39}$$

The branch cut in the logarithms must be taken as $-Q^2 \to -Q^2 - i\epsilon$. The hard function satisfies a multiplicative RGE, given by

$$\mu\frac{d}{d\mu} \ln H\left(Q^2, \mu\right) = 2\mathrm{Re}\left[\gamma_C\left(Q^2, \mu\right)\right] , \tag{A.40}$$

370

where $\gamma_C(Q^2, \mu)$ is the anomalous dimension for the Wilson coefficient, which is given to one-loop by

$$\gamma_C(Q^2, \mu) = \frac{\alpha_s C_F}{4\pi} \left(4\log\left[\frac{-Q^2}{\mu^2}\right] - 6 \right) . \tag{A.41}$$

Hard Splitting Function

The hard splitting function can be calculated using known results for the one-loop splitting functions [476] or from the result for $e^+e^- \to 3$ jets [477]. However, since at leading power the measurement of the 2-point energy correlation functions define the energy fractions and splitting angle, it is simplest to change variables in the results of Ref. [105], where the hard splitting function matching was performed for jet mass. Using the notation $t = s_{qg}$, and $x = s_{q\bar{q}}/Q^2$, Ref. [105] gave the matching coefficient to one-loop as

$$H_2^{q \to qg}(t, x, \mu) = Q^2 \frac{\alpha_s(\mu) C_F}{2\pi} \frac{1}{t} \frac{1 + x^2}{1 - x} \left\{ 1 + \frac{\alpha_s(\mu)}{2\pi} \left[\left(\frac{C_A}{2} - C_F \right) \left(2\log\frac{t}{\mu^2}\log x + \log^2 x + 2\mathrm{Li}_2(1 - x) \right) \right. \right.$$
$$\left. \left. - \frac{C_A}{2} \left(\log^2\frac{t}{\mu^2} - \frac{7\pi^2}{6} + 2\log\frac{t}{\mu^2}\log(1 - x) + \log^2(1 - x) + 2\mathrm{Li}_2(x) \right) + (C_A - C_F)\frac{1 - x}{1 + x^2} \right] \right\} . \tag{A.42}$$

We can now perform a change of variables to rewrite this in terms of $e_2^{(\alpha)}$, and the subjet energy fractions, using the leading power relation of Eq. (A.21), and the kinematic relations valid in the collinear subjets region of phase space. We find

$$t = \frac{Q^2}{2} \frac{(z_a z_b)^{1 - 2/\alpha} \left(e_2^{(\alpha)} \right)^{2/\alpha}}{2} , \qquad x = z_q , \tag{A.43}$$

and

$$H_2^{q \to qg}(e_2^{(\alpha)}, z_q, \mu) = \frac{\alpha_s(\mu) C_F}{\alpha\pi} \frac{1}{e_2^{(\alpha)}} \frac{1 + z_q^2}{1 - z_q} \tag{A.44}$$

$$\times \left\{ 1 + \frac{\alpha_s(\mu)}{2\pi} \left[\left(\frac{C_A}{2} - C_F \right) \left(2\log \left(\frac{Q^2}{\mu^2} \frac{(z_a z_b)^{1 - 2/\alpha} \left(e_2^{(\alpha)} \right)^{2/\alpha}}{4} \right) \log z_q + \log^2 z_q + 2\mathrm{Li}_2(1 - z_q) \right) \right. \right.$$

371

$$-\frac{C_A}{2}\left(\log^2\left(\frac{Q^2}{\mu^2}\frac{(z_a z_b)^{1-2/\alpha}\left(e_2^{(\alpha)}\right)^{2/\alpha}}{4}\right)-\frac{7\pi^2}{6}\right.$$

$$\left.+2\log\left(\frac{Q^2}{\mu^2}\frac{(z_a z_b)^{1-2/\alpha}\left(e_2^{(\alpha)}\right)^{2/\alpha}}{4}\right)\log(1-z_q)+\log^2(1-z_q)+2\mathrm{Li}_2(z_q)\right)+(C_A-C_F)\frac{1-z_q}{1+z_q^2}\right]\bigg\}\,.$$

Note that the hard splitting function depends on the partons involved in the split, which in our case we have taken to be $q \to qg$, and therefore singled out z_q, which is the energy fraction of the quark jet (defined identically to z_a, z_b). Throughout the rest of this appendix, we will, whenever possible, write results in terms of z_a, and z_b for generic partons, using general Casimirs. Since we consider the case $q \to qg$, we will calculate the jet functions for both quark and gluon jets, and therefore the results in this appendix are sufficient to treat general two-prong substructure, where the prongs are associated with generic partons by using the hard splitting function for other partonic splittings.

For completeness, we also present the one-loop results for $g \to gg$ and $g \to q\bar{q}$ splittings. While one-loop, and even two-loop, splitting helicity amplitudes exist in the literature [476, 478, 479], to our knowledge, the one-loop unpolarized splitting functions have not not been explicitly written down before. Using the results from Refs. [476, 479], the one-loop function for the $g \to gg$ splitting is

$$H_2^{g\to gg}(s_{gg},z,\mu)=\frac{\alpha_s(\mu)C_A}{2\pi}\frac{1}{s_{gg}}\left(\frac{z}{1-z}+\frac{1-z}{z}+z(1-z)\right)\left\{1+\frac{\alpha_s(\mu)}{2\pi}N\left[\log\frac{\mu^2}{s_{gg}}\log\left(z(1-z)\right)\right.\right.$$

$$\left.\left.-\frac{1}{2}\log^2\frac{s_{gg}}{\mu^2}-\frac{1}{2}\log^2\frac{z}{1-z}+\frac{5\pi^2}{12}+\left(\frac{1}{3}-\frac{n_F}{3N}\right)\frac{z(1-z)}{1+z^4+(1-z)^4}\right]\right\}\,. \qquad (A.45)$$

Here, we have expressed the result in terms of the numbers of colors, N, of the gauge theory and number of active quarks, n_F. Note that $C_A = N$. The virtuality of the splitting is

$$s_{gg}=2z_a z_b E_J^2(n_a \cdot n_b)\,, \qquad (A.46)$$

where a and b denote the final-state gluons in the splitting. Its anomalous dimension to

one-loop is

$$\gamma_{g \to gg} = \frac{\alpha_s(\mu)}{\pi} \left[N \log \frac{s_{gg}}{\mu^2} + N \log z(1-z) - \frac{\beta_0}{2} \right] . \tag{A.47}$$

For the one-loop result of the $g \to q\bar{q}$ splitting, we have

$$
\begin{aligned}
H_2^{g \to q\bar{q}}(s_{q\bar{q}}, z, \mu) = {} & \frac{\alpha_s(\mu) n_F}{2\pi} \frac{1}{s_{q\bar{q}}} \left(z^2 + (1-z)^2 \right) \left\{ 1 + \frac{\alpha_s(\mu)}{2\pi} \left[N \log \frac{\mu^2}{s_{q\bar{q}}} \log(z(1-z)) \right. \right. \\
& + \frac{3}{2} \frac{1}{N} \log \frac{\mu^2}{s_{q\bar{q}}} - \frac{2n_F}{3} \log \frac{\mu^2}{s_{q\bar{q}}} + \frac{13}{6} N \log \frac{\mu^2}{s_{q\bar{q}}} + \frac{1}{2N} \log^2 \frac{\mu^2}{s_{q\bar{q}}} \\
& \left. \left. - \frac{1}{N} \frac{7\pi^2}{12} - N \frac{\pi^2}{6} - \frac{N}{2} \log^2 \frac{z}{1-z} + \frac{40}{9} N - \frac{10}{9} n_F \right] \right\} .
\end{aligned}
\tag{A.48}
$$

Note that, in terms of the number of colors,

$$C_F = \frac{N^2 - 1}{2N} .$$

Its anomalous dimension is

$$\gamma_{g \to q\bar{q}} = \frac{\alpha_s(\mu)}{\pi} \left[\frac{1}{N} \log \frac{\mu^2}{s_{q\bar{q}}} + N \log(z(1-z)) + \frac{\beta_0}{2} - 3C_F \right] . \tag{A.49}$$

Global Soft Function

In this section we calculate the global soft function. The global soft modes can resolve the boundaries of the jet, so the jet algorithm constraint cannot be expanded. However, the soft modes do not resolve the dipole of the collinear splitting. The global soft function therefore has two Wilson lines in the n and $\bar{n}$ directions. A general one-loop soft function can be written as

$$S_G^{(1)} \left(e_3^{(\alpha)} \right) = \frac{1}{2} \sum_{i \neq j} \mathbf{T}_i \cdot \mathbf{T}_j \, S_{G,ij}^{(1)} \left(e_3^{(\alpha)} \right) , \tag{A.50}$$

373

where $\mathbf{T}_i$ is the color generator of leg i in the notation of Refs. [388, 389], and the sum runs over all pairs of legs. Here we have only the contribution from $i, j = n, \bar{n}$, but we still perform this extraction of the color structure to keep the results generic.

The one-loop integrand for the soft function is given by

$$S_{G,n\bar{n}}^{(1)}(e_3^{(\alpha)}) = \tag{A.51}$$
$$-g^2 \left(\frac{\mu^2 e^{\gamma_E}}{4\pi}\right)^\epsilon \int [d^d k]_+ \frac{2n \cdot \bar{n}}{n \cdot k \, k \cdot \bar{n}} \Theta\left(\tan^2\frac{R}{2} - \frac{n \cdot k}{\bar{n} \cdot k}\right) \delta\left(e_3^{(\alpha)} - N_S \left(\frac{\bar{n} \cdot k + n \cdot k}{2Q}\right)^{1-\alpha} \left(\frac{n \cdot k}{Q}\right)^\alpha\right)$$

with $d = 4 - 2\epsilon$, and where here we have extracted the normalization factor

$$N_S = 2^{3\alpha/2+1} z_a z_b \, (n_a \cdot n_b)^{\alpha/2} \,, \tag{A.52}$$

following the expression for the three point energy correlation function in the soft power counting, given in Eq. (A.27). The first Θ-function in Eq. (A.51) implements the jet algorithm constraint, which is simple for a single emission. To simplify notation, we also use the following shorthand for the measure for a positive energy, on-shell, collinear particle

$$[d^d k]_+ = \frac{d^d k}{(2\pi)^d} 2\pi \Theta(\bar{n} \cdot k) \delta(k^2) \,. \tag{A.53}$$

To perform this integral, it is convenient to make the change of variables

$$\bar{n} \cdot k = v \,, \qquad n \cdot k = v \, u \,, \tag{A.54}$$

which factorizes the jet algorithm constraint and the measurement function. The integrals can then be evaluated using standard techniques. Performing all the integrals but the u integral, and transforming to Laplace space, $e_3^{(\alpha)} \to \tilde{e}_3^{(\alpha)}$, gives

$$\tilde{S}_{G,n\bar{n}}^{(1)}(\tilde{e}_3^{(\alpha)}) = -\frac{g^2 e^{-\epsilon\gamma_E}\Gamma(-2\epsilon)}{(2\pi)^2\Gamma(1-\epsilon)} \left(\frac{e^{\gamma_E}\mu \tilde{e}_3^{(\alpha)} N_S}{2^{1-\alpha}Q}\right)^{2\epsilon} \int_0^{\tan^2\frac{R}{2}} \frac{du}{u^{1+\epsilon(1-2\alpha)}}(1+u)^{2\epsilon(1-\alpha)} \,. \tag{A.55}$$

This can be integrated exactly in terms of hypergeometric functions,

$$\int_0^{\tan^2\frac{R}{2}} \frac{du}{u^{1+\epsilon(1-2\alpha)}}(1+u)^{2\epsilon(1-\alpha)} = \frac{\Gamma(-\epsilon(1-2\alpha))}{\Gamma(1-\epsilon(1-2\alpha))}$$

$$\times \left(\tan^2\frac{R}{2}\right)^\epsilon \left(\frac{\tan^2\frac{R}{2}}{1+\tan^2\frac{R}{2}}\right)^{-2(1-\alpha)\epsilon} {}_2F_1\left[1,-2(1-\alpha)\epsilon;1-(1-2\alpha)\epsilon;\frac{\tan^2\frac{R}{2}}{1+\tan^2\frac{R}{2}}\right], \tag{A.56}$$

where we have used both a Pfaff and an Euler transformation to extract the singular behavior from the hypergeometric function. We therefore have

$$\tilde{S}_{G,n\bar{n}}^{(1)}(\tilde{e}_3^{(\alpha)}) = -\frac{\alpha_s}{\pi}\frac{e^{-\epsilon\gamma_E}\Gamma(-2\epsilon)}{\Gamma(1-\epsilon)}\left(\frac{e^{\gamma_E}\mu\tilde{e}_3^{(\alpha)}N_S}{2^{1-\alpha}Q}\tan\frac{R}{2}\right)^{2\epsilon}\frac{\Gamma(-\epsilon(1-2\alpha))}{\Gamma(1-\epsilon(1-2\alpha))} \tag{A.57}$$

$$\times \left(\frac{\tan^2\frac{R}{2}}{1+\tan^2\frac{R}{2}}\right)^{-2(1-\alpha)\epsilon} {}_2F_1\left[1,-2(1-\alpha)\epsilon;1-(1-2\alpha)\epsilon;\frac{\tan^2\frac{R}{2}}{1+\tan^2\frac{R}{2}}\right].$$

Expanding in ϵ (throughout these appendices we use the HypExp package [480, 481] for expansions of hypergeometric functions) and separating in divergent and finite pieces, we find

$$\tilde{S}_{G,n\bar{n}}^{(1)\text{div}}(\tilde{e}_3^{(\alpha)}) = \frac{\alpha_s}{2\pi}\frac{1}{(2\alpha-1)\epsilon^2} + \frac{\alpha_s}{\pi}\frac{\log\left[\frac{e^{\gamma_E}\mu\tilde{e}_3^{(\alpha)}N_S}{2^{1-\alpha}Q}\right]}{(2\alpha-1)\epsilon} + \frac{\alpha_s}{2\pi}\frac{\log\left[\tan^2\frac{R}{2}\right]}{\epsilon}, \tag{A.58}$$

$$\tilde{S}_{G,n\bar{n}}^{(1)\text{fin}}(\tilde{e}_3^{(\alpha)}) = \frac{\alpha_s}{\pi}\left\{\frac{\log^2\left[\frac{e^{\gamma_E}\mu\tilde{e}_3^{(\alpha)}N_S}{2^{1-\alpha}Q}\right]}{2\alpha-1} + \log\left[\frac{e^{\gamma_E}\mu\tilde{e}_3^{(\alpha)}N_S}{2^{1-\alpha}Q}\right]\log\left[\tan^2\frac{R}{2}\right] \right. \tag{A.59}$$

$$\left. + \frac{\pi^2}{8(2\alpha-1)} + \frac{2\alpha-1}{4}\log^2\left[\tan^2\frac{R}{2}\right] + (\alpha-1)\text{Li}_2\left[-\tan^2\frac{R}{2}\right]\right\},$$

where Li_2 is the dilogarithm function.

Jet Function

To calculate the jet function, we use the approach of Ref. [330] and integrate the appropriate splitting functions against our measurement function. In the power counting of the jet function, we can expand the jet algorithm constraint

$$\Theta\left(\tan^2\frac{R}{2} - \frac{n \cdot k}{\bar{n} \cdot k}\right) \to 1. \tag{A.60}$$

The one-loop jet function in the n_a direction is then given by

$$J_{i,n_a}^{(1)}(Q_J, e_3^{(\alpha)}) = \int d\Phi_2^c \, \sigma_2^c \, \delta\left(e_3^{(\alpha)} - N_J \left(\frac{\bar{n}_a \cdot k_1}{Q}\right)^{1-\alpha/2} \left(\frac{\bar{n}_a \cdot k_2}{Q}\right)^{1-\alpha/2} \left(\frac{k_1 \cdot k_2}{Q^2}\right)^{\alpha/2}\right). \tag{A.61}$$

The two particle collinear phase space is given by [232]

$$d\Phi_2^c = 2(2\pi)^{3-2\epsilon} Q_J \left[d^d k_1\right]_+ \left[d^d k_2\right]_+ \delta(Q_J - \bar{n}_a \cdot k_1 - \bar{n}_a \cdot k_2) \delta^{d-2}(k_{1\perp} + k_{2\perp}), \tag{A.62}$$

and

$$\sigma_2^c = \left(\frac{\mu^2 e^{\gamma_E}}{4\pi}\right)^\epsilon \frac{2g^2}{s} P_i(z), \tag{A.63}$$

where

$$P_q(z) = C_F \left[\frac{1+z^2}{1-z} - \epsilon(1-z)\right], \tag{A.64}$$

and

$$P_g = C_A \left[\frac{z}{1-z} + \frac{1-z}{z} + z(1-z)\right] + \frac{n_f}{2}\left[1 - \frac{2z(1-z)}{1-\epsilon}\right], \tag{A.65}$$

which includes both the $g \to gg$ and $g \to q\bar{q}$ contributions. Explicitly, the integrand is then given by

$$J_{i,n_a}^{(1)}(Q_J, e_3^{(\alpha)}) = \left(\frac{\mu^2 e^{\gamma_E}}{4\pi}\right)^\epsilon 2(2\pi)^{3-2\epsilon} Q_J 2g^2 \int [d^d k_1]_+ \int [d^d k_2]_+ \frac{P_i\left(\frac{\bar{n}_a \cdot k_1}{Q_J}\right)}{2k_1 \cdot k_2} \tag{A.66}$$

$$\times \, \delta(Q_J - \bar{n}_a \cdot k_1 - \bar{n}_a \cdot k_2) \delta^{d-2}(\vec{k}_{1\perp} + \vec{k}_{2\perp})$$

376

$$\times \delta\left(e_3^{(\alpha)} - N_J \left(\frac{\bar{n}_a \cdot k_1}{Q}\right)^{1-\alpha/2} \left(\frac{\bar{n}_a \cdot k_2}{Q}\right)^{1-\alpha/2} \left(\frac{k_1 \cdot k_2}{Q^2}\right)^{\alpha/2} \right),$$

where we have extracted the normalization factor

$$N_J = 2^{5\alpha/2}(n_a \cdot n_b)^\alpha z_b, \tag{A.67}$$

for simplicity, following the expression of Eq. (A.24) for the three point energy correlation function in the power counting for the emission of a single collinear particle. Furthermore, note that we have used $Q_J = z_a Q$ in this expression.

The integrals can be performed using standard techniques, and we find, after transforming to Laplace space, $e_3^{(\alpha)} \to \tilde{e}_3^{(\alpha)}$, for the jet function in the n_a direction

$$\tilde{J}_{g,n_a}^{(1)}(Q_J, \tilde{e}_3^{(\alpha)}) = \frac{\alpha_s}{2\pi} C_A \left(\frac{\alpha}{(\alpha-1)\epsilon^2} + \frac{2L_\alpha^{J,a}\left(\tilde{e}_3^{(\alpha)}\right)}{(\alpha-1)\epsilon} + \frac{1}{\epsilon} \frac{11C_A - 2n_f}{6C_A} \right. \tag{A.68}$$

$$- \frac{\alpha\pi^2}{12(\alpha-1)} + \frac{\pi^2}{3(\alpha-1)\alpha} - \frac{67}{9\alpha} + \frac{2\pi^2}{3\alpha} + \frac{2L_\alpha^{J,a}\left(\tilde{e}_3^{(\alpha)}\right)^2}{(\alpha-1)\alpha} + \frac{11L_\alpha^{J,a}\left(\tilde{e}_3^{(\alpha)}\right)}{3\alpha}$$

$$\left. + \frac{67}{9} - \frac{2\pi^2}{3} - \frac{2n_f L_\alpha^{J,a}\left(\tilde{e}_3^{(\alpha)}\right)}{3C_A\alpha} + \frac{13n_f}{9C_A\alpha} - \frac{23n_f}{18C_A} \right),$$

for gluon jets, and

$$\tilde{J}_{q,n_a}^{(1)}(Q_J, \tilde{e}_3^{(\alpha)}) = \frac{\alpha_s}{2\pi} C_F \left(-\frac{\alpha}{\epsilon^2(1-\alpha)} + \frac{3}{2\epsilon} - \frac{2}{\epsilon(1-\alpha)} L_\alpha^{J,a}\left(\tilde{e}_3^{(\alpha)}\right) \right. \tag{A.69}$$

$$- \frac{2}{\alpha(1-\alpha)} L_\alpha^{J,a}\left(\tilde{e}_3^{(\alpha)}\right)^2 + \frac{3}{\alpha} L_\alpha^{J,a}\left(\tilde{e}_3^{(\alpha)}\right) - \frac{\pi^2}{6\alpha}$$

$$\left. - \frac{\pi^2}{4\alpha(1-\alpha)} + \frac{3\pi^2(1-\alpha)}{4\alpha} + \frac{1}{2\alpha} - \frac{13(1-\alpha)}{2\alpha} \right),$$

for quark jets respectively. The jet function for the n_b direction can be trivially found from $a \to b$.

Here we have used $L_\alpha^{J,a}\left(\tilde{e}_3^{(\alpha)}\right)$ to denote the logarithm appearing in the jet functions. The

argument of this logarithm depends on the subjet energy fraction. We indicate the specific logarithm for the subjet via the notation

$$L_\alpha^{J,a}\left(\tilde{e}_3^{(\alpha)}\right) = \log\left[N_J\tilde{e}_3^{(\alpha)}e^{\gamma_E}\left(\frac{\mu}{\sqrt{2}Q}\right)^\alpha z_a^{2-\alpha}\right].$$ (A.70)

Collinear-Soft Function

We now calculate the collinear-soft function. The collinear-soft modes couple eikonally to the collinear sector, and so the collinear-soft function has the one-loop form

$$S_c^{(1)}\left(e_3^{(\alpha)}\right) = \frac{1}{2}\sum_{i\neq j}\mathbf{T}_i\cdot\mathbf{T}_j\,S_{c,ij}^{(1)}\left(e_3^{(\alpha)}\right),$$ (A.71)

where $\mathbf{T}_i$ is the color generator of leg i in the notation of Refs. [388, 389], and the sum runs over all pairs of legs. Since the collinear-soft modes resolves the dipole from the collinear splitting, there are three Wilson lines, $n_a, n_b, \bar{n}$ to which the collinear-soft modes couple. We calculate separately the contributions arising from the pair of legs n_a, n_b, and from the pairs $n_{a,b}, \bar{n}$. In both cases the integral involves the jet algorithm constraint. In the power counting of the collinear-soft modes, this constraint can be expanded as

$$\Theta\left(\tan^2\frac{R}{2} - \frac{n\cdot k}{\bar{n}\cdot k}\right)\to 1.$$ (A.72)

If this expansion was not performed, the contribution of the collinear soft modes sensitive to the jet radius R, would be removed by a soft zero bin subtraction.

n_a, n_b Contribution:

We begin by calculating the contribution from the emission between the n_a, n_b eikonal lines. The integrand is given by

$$S_{c,n_a n_b}^{(1)}(e_3^{(\alpha)}) =$$ (A.73)

378

$$-g^2\left(\frac{\mu^2 e^{\gamma_E}}{4\pi}\right)^\epsilon\int [d^dk]_+\frac{2n_a\cdot n_b}{n_a\cdot k\,k\cdot n_b}\delta\left(e_3^{(\alpha)}-N_{CS}\left(\frac{\bar{n}\cdot k}{2Q}\right)^{1-\alpha}\left(\frac{n_a\cdot k}{Q}\right)^{\alpha/2}\left(\frac{n_b\cdot k}{Q}\right)^{\alpha/2}\right),$$

where we have extracted the normalization factor

$$N_{CS}=2^{3\alpha/2+1}z_a z_b\left(n_a\cdot n_b\right)^{\alpha/2}, \tag{A.74}$$

for simplicity, following the expression of Eq. (A.28) for the three point energy correlation function in the power counting for the emission of a single collinear-soft particle.

To perform the calculation, we go to the light-cone basis defined by $n,\bar{n}$. We then have

$$
\begin{aligned}
n_a\cdot k &= \frac{n\cdot n_a}{2}\bar{n}\cdot k+\frac{\bar{n}\cdot n_a}{2}n\cdot k+k_\perp\cdot n_{a\perp}\\
&= \frac{n\cdot n_a}{2}\bar{n}\cdot k+\frac{\bar{n}\cdot n_a}{2}n\cdot k-(\bar{n}\cdot k\,n\cdot k)^{1/2}|\hat{n}_{a\perp}|\cos\theta, \tag{A.75}\\
n_b\cdot k &= \frac{n\cdot n_b}{2}\bar{n}\cdot k+\frac{\bar{n}\cdot n_b}{2}n\cdot k+k_\perp\cdot n_{b\perp}\\
&= \frac{n\cdot n_b}{2}\bar{n}\cdot k+\frac{\bar{n}\cdot n_b}{2}n\cdot k+(\bar{n}\cdot k\,n\cdot k)^{1/2}|\hat{n}_{a\perp}|\cos\theta, \tag{A.76}
\end{aligned}
$$

where θ denotes the angle between the particle k and the n axis. In the above kinematic relations, we have made use of the fact that since $\hat{n}\sim\hat{n}_a+\hat{n}_b$, $k_\perp\cdot n_{b\perp}=-k_\perp\cdot n_{a\perp}$. Rewriting the integrand for a positive energy gluon in terms of θ, we find

$$\int [d^dk]_+ = \frac{1}{2^{4-2\epsilon}\pi^{\frac{5}{2}-\epsilon}\Gamma(\frac{1}{2}-\epsilon)}\int_0^\infty\frac{dn\cdot k}{n\cdot k^\epsilon}\int_0^\infty\frac{d\bar{n}\cdot k}{\bar{n}\cdot k^\epsilon}\int_0^\pi d\theta\sin^{-2\epsilon}\theta, \tag{A.77}$$

$$= c_\epsilon\int_0^\infty\frac{dn\cdot k}{n\cdot k^\epsilon}\int_0^\infty\frac{d\bar{n}\cdot k}{\bar{n}\cdot k^\epsilon}\int_0^\pi d\theta\sin^{-2\epsilon}\theta, \tag{A.78}$$

for $d=4-2\epsilon$. To simplify our expressions, we have extracted the following constant

$$c_\epsilon=\frac{1}{2^{4-2\epsilon}\pi^{\frac{5}{2}-\epsilon}\Gamma(\frac{1}{2}-\epsilon)}. \tag{A.79}$$

In the collinear soft region of phase space, we power count $n_a\cdot n_b\ll 1$. We can therefore work to leading power in $n_a\cdot n_b$ in the integrand. Using the relations of Eq. (A.15)- Eq. (A.17),

and expanding to leading power in $n_a \cdot n_b$, we have

$$n_a \cdot k \, n_b \cdot k = \left(n \cdot k + \frac{n_a \cdot n_b}{8} \bar{n} \cdot k\right)^2 - \frac{n_a \cdot n_b}{2}(n \cdot k \, \bar{n} \cdot k)\cos^2\theta. \qquad (A.80)$$

Note that in our power counting, $n \cdot k \sim n_a \cdot n_b$, so that this expression scales homogeneously. To perform the integral, we make the change of variables

$$\bar{n} \cdot k = v, \qquad n \cdot k = vw\left(\frac{n_a \cdot n_b}{8}\right). \qquad (A.81)$$

We then have

$$n_a \cdot k \, n_b \cdot k = v^2 \left(\frac{n_a \cdot n_b}{8}\right)^2 \left[(1+w)^2 - 4w\cos^2\theta\right] \qquad (A.82)$$

$$= v^2 \left(\frac{n_a \cdot n_b}{8}\right)^2 \left[(1-w)^2 + 4w\sin^2\theta\right]. \qquad (A.83)$$

The one loop expression for the collinear soft function can then be written

$$S^{(1)}_{c,\,n_a n_b}(e_3^{(\alpha)}) = \qquad (A.84)$$

$$-g^2 \left(\frac{\mu^2 e^{\gamma_E}}{4\pi}\right)^\epsilon 16c_\epsilon \left(\frac{n_a \cdot n_b}{8}\right)^{-\epsilon} \int_0^\infty \frac{dw}{w^\epsilon} \int_0^\infty \frac{dv}{v^{1+2\epsilon}} \int_0^\pi d\theta \sin^{-2\epsilon}\theta \frac{1}{(1-w)^2 + 4w\sin^2\theta}$$

$$\times \delta\left(e_3^{(\alpha)} - \frac{N_{CS}}{2^{1-\alpha}} \frac{v}{Q} \left(\frac{n_a \cdot n_b}{8}\right)^\alpha \left[(1-w)^2 + 4w\sin^2\theta\right]^{\alpha/2}\right),$$

The v integral is straightforward. Transforming to Laplace space, $e_3^{(\alpha)} \to \tilde{e}_3^{(\alpha)}$, we find

$$S^{(1)}_{c,\,n_a n_b}(\tilde{e}_3^{(\alpha)}) = -g^2 \Gamma(-2\epsilon) \left(\frac{\mu^2 N_{CS}^2 e^{\gamma_E}(\tilde{e}_3^{(\alpha)})^2 \left(\frac{n_a \cdot n_b}{8}\right)^{-1+2\alpha}}{4\pi 4^{1-\alpha} Q^2}\right)^\epsilon \qquad (A.85)$$

$$\times 16c_\epsilon \int_0^\infty \frac{dw}{w^\epsilon} \int_0^\pi d\theta \sin^{-2\epsilon}\theta \left[(1-w)^2 + 4w\sin^2\theta\right]^{-1+\alpha\epsilon}.$$

The θ integral can be performed exactly in terms of hypergeometric functions using

$$\int_0^\pi d\theta \sin^{-2\epsilon}\theta \left[(1-w)^2 + 4w\sin^2\theta\right]^{-1+\alpha\epsilon} = \qquad (A.86)$$

$$\frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]}(1-w)^{2(-1+\alpha\epsilon)}\,{}_2F_1\left[1-\alpha\epsilon,1/2-\epsilon,1-\epsilon,-\frac{4w}{(1-w)^2}\right],$$

which can be rewritten using a Pfaff transformation as

$$\int_0^\pi d\theta \sin^{-2\epsilon}\theta\left[(1-w)^2+4w\sin^2\theta\right]^{-1+\alpha\epsilon}=\frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]} \tag{A.87}$$

$$\times\,(1+w)^{-1+2\epsilon}\left((1-w)^2\right)^{-1/2-(1-\alpha)\epsilon}\,{}_2F_1\left[1/2-\epsilon,-\epsilon+\alpha\epsilon,1-\epsilon,\frac{4w}{(1+w)^2}\right].$$

The remaining integral in w is given by

$$S^{(1)}_{c,n_an_b}(\tilde{e}_3^{(\alpha)})=-g^2\Gamma(-2\epsilon)\left(\frac{\mu^2N_{CS}^2e^{\gamma_E}(\tilde{e}_3^{(\alpha)})^2\left(\frac{n_a\cdot n_b}{8}\right)^{-1+2\alpha}}{4\pi 4^{1-\alpha}Q^2}\right)^\epsilon 16c_\epsilon\frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]}$$

$$\int_0^\infty \frac{dw}{w^\epsilon}(1+w)^{-1+2\epsilon}\left((1-w)^2\right)^{-1/2-(1-\alpha)\epsilon}\,{}_2F_1\left[1/2-\epsilon,-\epsilon+\alpha\epsilon,1-\epsilon,\frac{4w}{(1+w)^2}\right].$$

$$\tag{A.88}$$

Re-mapping the integral to the unit interval, we have

$$S^{(1)}_{c,n_an_b}(\tilde{e}_3^{(\alpha)})=-g^2\Gamma(-2\epsilon)\left(\frac{\mu^2N_{CS}^2e^{\gamma_E}(\tilde{e}_3^{(\alpha)})^2\left(\frac{n_a\cdot n_b}{8}\right)^{-1+2\alpha}}{4\pi 4^{1-\alpha}Q^2}\right)^\epsilon 16c_\epsilon\frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]}$$

$$\int_0^1 dw\left(w^{-\epsilon}+w^{(1-2\alpha)\epsilon}\right)(1+w)^{-1+2\epsilon}(1-w)^{-1-2(1-\alpha)\epsilon}\,{}_2F_1\left[1/2-\epsilon,-\epsilon+\alpha\epsilon,1-\epsilon,\frac{4w}{(1+w)^2}\right].$$

We could not perform this integral exactly, but it can be done as a Laurent expansion in ϵ by expanding the hypergeometric function as

$$_2F_1\left[\frac{1}{2}-\epsilon,-(1-\alpha)\epsilon;1-\epsilon;\frac{4w}{(1+w)^2}\right]=1-2(1-\alpha)\epsilon\ln(1+w)+\mathcal{O}(\epsilon^2), \tag{A.89}$$

which is valid for $0\leq w\leq 1$, and we have truncated the expansion at $\mathcal{O}(\epsilon^2)$ as we are only interested in the terms up to $\mathcal{O}(\epsilon^0)$ in the one-loop result. We then have

$$S^{(1)}_{c,n_an_b}(\tilde{e}_3^{(\alpha)})=-g^2\Gamma(-2\epsilon)\left(\frac{\mu^2N_{CS}^2e^{\gamma_E}(\tilde{e}_3^{(\alpha)})^2\left(\frac{n_a\cdot n_b}{8}\right)^{-1+2\alpha}}{4\pi 4^{1-\alpha}Q^2}\right)^\epsilon 16c_\epsilon\frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]}$$

$$\int_0^1 dw \left(w^{-\epsilon} + w^{(1-2\alpha)\epsilon} \right)(1+w)^{-1+2\epsilon}(1-w)^{-1-2(1-\alpha)\epsilon}\left(1 - 2(1-\alpha)\epsilon \ln(1+w)\right).$$

<div align="right">(A.90)</div>

For the remaining integral in w, we have

$$\int_0^1 dw \left(w^{-\epsilon} + w^{(1-2\alpha)\epsilon} \right)(1+w)^{-1+2\epsilon}(1-w)^{-1-2(1-\alpha)\epsilon}\left(1 - 2(1-\alpha)\epsilon \ln(1+w)\right) =$$

$$\int_0^1 dw \left(w^{-\epsilon} + w^{(1-2\alpha)\epsilon} \right)(1+w)^{-1+2\epsilon}(1-w)^{-1-2(1-\alpha)\epsilon}$$

<div align="right">(A.91)</div>

$$- 2(1-\alpha)\epsilon \int_0^1 dw \left(w^{-\epsilon} + w^{(1-2\alpha)\epsilon} \right)(1+w)^{-1+2\epsilon}(1-w)^{-1-2(1-\alpha)\epsilon} \log(1+w).$$

The first integral can be done in terms of hypergeometric functions, while the second can be done using plus functions (for a detailed discussion of their properties, see e.g. [195]), and applying the identity

$$\frac{1}{z^{1+a\epsilon}} = -\frac{1}{a\epsilon}\delta(z) + \sum_{i=0}^{\infty} \frac{(-a\epsilon)^i}{i!}\mathcal{D}_i(z),$$

<div align="right">(A.92)</div>

with

$$\mathcal{D}_i(z) = \left[\frac{\log^i z}{z} \right]_+ .$$

<div align="right">(A.93)</div>

We find

$$\int_0^1 dw \left(w^{-\epsilon} + w^{(1-2\alpha)\epsilon} \right)\left((1+w)^2\right)^{-\frac{1}{2}+\epsilon}\left((1-w)^2\right)^{-\frac{1}{2}-(1-\alpha)\epsilon}\left(1 - 2(1-\alpha)\epsilon \ln(1+w)\right)$$

$$= \frac{\Gamma[2(\alpha-1)\epsilon]\Gamma[1-\epsilon]}{\Gamma[1-3\epsilon+2\alpha\epsilon]}{}_2F_1[1-2\epsilon, 1-\epsilon; 1-3\epsilon+2\alpha\epsilon; -1]$$

<div align="right">(A.94)</div>

$$+ \frac{\Gamma[2(\alpha-1)\epsilon]\Gamma[1+\epsilon-2\alpha\epsilon]}{\Gamma[1-\epsilon]}{}_2F_1[1-2\epsilon, 1+\epsilon-2\alpha\epsilon; 1-\epsilon; -1]$$

$$+ 2^{2\epsilon}\log 2 - 2(1-\alpha)\epsilon \left(\log^2 2 - \frac{\pi^2}{12} \right)$$

$$= \frac{1}{(2\alpha-2)\epsilon} + \frac{\alpha\log(2)}{\alpha-1} + \log(2) + \frac{\epsilon\left(-\pi^2\alpha^2 + 36\alpha^2\log^2(2) + 3\pi^2\alpha - 24\alpha\log^2(2) - 2\pi^2\right)}{12(\alpha-1)}.$$

Therefore, in total, we have

$$S_{c,n_a n_b}^{(1)}(\tilde{e}_3^{(\alpha)}) = -g^2\Gamma(-2\epsilon)\left(\frac{\mu^2 N_{CS}^2 e^{\gamma_E}(\tilde{e}_3^{(\alpha)})^2\left(\frac{n_a\cdot n_b}{8}\right)^{-1+2\alpha}}{4\pi 4^{1-\alpha}Q^2}\right)^\epsilon 16c_\epsilon\frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]}$$

$$\left(\frac{1}{(2\alpha-2)\epsilon} + \frac{\alpha\log(2)}{\alpha-1} + \log(2) + \frac{\epsilon\left(-\pi^2\alpha^2 + 36\alpha^2\log^2(2) + 3\pi^2\alpha - 24\alpha\log^2(2) - 2\pi^2\right)}{12(\alpha-1)}\right).$$

(A.95)

Expanding in ϵ, and keeping only the divergent piece, as relevant for the anomalous dimensions, we find

$$\tilde{S}_{c,n_a n_b}^{(1)\text{div}}(\tilde{e}_3^{(\alpha)}) = \frac{\alpha_s}{\pi}\frac{1}{(\alpha-1)\epsilon^2} + 2\frac{\alpha_s}{\pi}\frac{\left(2\alpha\log(2) + \log\left[\frac{\mu N_{CS}e^{\gamma_E}(\tilde{e}_3^{(\alpha)})\left(\frac{n_a\cdot n_b}{8}\right)^{-1/2+\alpha}}{2^{1-\alpha}Q}\right] - \log(2)\right)}{(\alpha-1)\epsilon}$$

$$= \frac{\alpha_s}{\pi}\frac{1}{(\alpha-1)\epsilon^2} + 2\frac{\alpha_s}{\pi}\frac{L_\alpha^{cs}}{(\alpha-1)\epsilon},$$

(A.96)

where

$$L_\alpha^{cs} = \log\left(\frac{\mu N_{CS}e^{\gamma_E}(\tilde{e}_3^{(\alpha)})(n_a\cdot n_b)^{-1/2+\alpha}}{\sqrt{2}Q}\right).$$

(A.97)

$n_a, \bar{n}$ and $n_b, \bar{n}$ **Contributions:**

We now calculate the $n_a, \bar{n}$ contribution to the collinear-soft function. The $n_b, \bar{n}$ contribution will be identical. The one-loop integrand is given by

$$S_{c,n_a \bar{n}}^{(1)}(e_3^{(\alpha)}) = \tag{A.98}$$

$$-g^2\left(\frac{\mu^2 e^{\gamma_E}}{4\pi}\right)^\epsilon\int[d^d k]_+\frac{2n_a\cdot\bar{n}}{n_a\cdot k\, k\cdot\bar{n}}\delta\left(e_3^{(\alpha)} - N_{CS}\left(\frac{\bar{n}\cdot k}{2Q}\right)^{1-\alpha}\left(\frac{n_a\cdot k}{Q}\right)^{\alpha/2}\left(\frac{n_b\cdot k}{Q}\right)^{\alpha/2}\right),$$

where we have again extracted the normalization factor

$$N_{CS} = 2^{3\alpha/2+1}z_a z_b\,(n_a\cdot n_b)^{\frac{\alpha}{2}}.$$

(A.99)

As with the $n_a \cdot n_b$ contribution, we expand the integrand to leading power in $n_a \cdot n_b$ using

$$n_a \cdot k \, n_b \cdot k = \left(n \cdot k + \frac{n_a \cdot n_b}{8} \bar{n} \cdot k\right)^2 - \frac{n_a \cdot n_b}{2}(n \cdot k \, \bar{n} \cdot k) \cos^2\theta \,, \tag{A.100}$$

$$n_a \cdot \bar{n} = 2 \,, \tag{A.101}$$

$$n_a \cdot k = \frac{n_a \cdot n_b}{8} \bar{n} \cdot k + n \cdot k - (n \cdot k \bar{n} \cdot k)^{1/2} \sqrt{\frac{n_a \cdot n_b}{2}} \cos\theta \,. \tag{A.102}$$

To perform the integral, it is again convenient to make the change of variables

$$\bar{n} \cdot k = v, \qquad n \cdot k = vw \left(\frac{n_a \cdot n_b}{8}\right) \,. \tag{A.103}$$

We then have

$$n_a \cdot k \, n_b \cdot k = v^2 \left(\frac{n_a \cdot n_b}{8}\right)^2 \left[(1-w)^2 + 4w \sin^2\theta\right] \,, \tag{A.104}$$

$$n_a \cdot k = \frac{n_a \cdot n_b}{8} v + vw \left(\frac{n_a \cdot n_b}{8}\right) - \left(v^2 w \left(\frac{n_a \cdot n_b}{8}\right)\right)^{1/2} \sqrt{\frac{n_a \cdot n_b}{2}} \cos\theta$$

$$= v \left(\frac{n_a \cdot n_b}{8}\right) \left(1 + w - 2\sqrt{w} \cos\theta\right) \,. \tag{A.105}$$

The one-loop expression for the contribution to the collinear soft function can then be written

$$S^{(1)}_{c, \, n_a \bar{n}}(e_3^{(\alpha)}) = \tag{A.106}$$

$$- g^2 \left(\frac{\mu^2 e^{\gamma_E}}{4\pi}\right)^\epsilon 4 c_\epsilon \left(\frac{n_a \cdot n_b}{8}\right)^{-\epsilon} \int_0^\infty \frac{dw}{w^\epsilon} \int_0^\infty \frac{dv}{v^{1+2\epsilon}} \int_0^\pi d\theta \sin^{-2\epsilon}\theta \frac{1}{1 + w - 2\sqrt{w} \cos\theta}$$

$$\delta \left(e_3^{(\alpha)} - \frac{N_{CS}}{2^{1-\alpha}} \frac{v}{Q} \left(\frac{n_a \cdot n_b}{8}\right)^\alpha \left[(1-w)^2 + 4w \sin^2\theta\right]^{\alpha/2}\right) \,.$$

This integral can be performed in a similar manner to the $n_a \cdot n_b$ integral. The v integral is straightforward, after transforming to Laplace space $e_3^{(\alpha)} \to \tilde{e}_3^{(\alpha)}$, we find

$$S^{(1)}_{c, \, n_a \bar{n}}(\tilde{e}_3^{(\alpha)}) = -g^2 4 c_\epsilon \Gamma(-2\epsilon) \left(\frac{\mu^2 N_{CS}^2 e^{\gamma_E} (\tilde{e}_3^{(\alpha)})^2 \left(\frac{n_a \cdot n_b}{8}\right)^{-1+2\alpha}}{4\pi 4^{1-\alpha} Q^2}\right)^\epsilon$$

384

$$\int_0^\infty \frac{dw}{w^\epsilon} \int_0^\pi d\theta \, \sin^{-2\epsilon}\theta \frac{\left[(1-w)^2 + 4w \, \sin^2\theta\right]^{\alpha\epsilon}}{1 + w - 2\sqrt{w}\cos\theta} \, . \tag{A.107}$$

We now focus on the integral

$$\int_0^\infty \frac{dw}{w^\epsilon} \int_0^\pi d\theta \, \sin^{-2\epsilon}\theta \frac{\left[(1-w)^2 + 4w \, \sin^2\theta\right]^{\alpha\epsilon}}{1 + w - 2\sqrt{w}\cos\theta} \, . \tag{A.108}$$

Remapping to the unit interval, we find

$$\int_0^1 du \, \left[u^{-\epsilon} + u^{-1+(1-2\alpha)\epsilon}\right] \int_0^\pi d\theta \, \sin^{-2\epsilon}\theta \frac{\left[(1-u)^2 + 4u \, \sin^2\theta\right]^{\alpha\epsilon}}{1 + u - 2\sqrt{u}\cos\theta}$$

$$= \int_0^1 du \, \left[u^{-\epsilon} + u^{-1+(1-2\alpha)\epsilon}\right]$$

$$\times \int_0^\pi d\theta \, \sin^{-2\epsilon}\theta \left[(1-u)^2 + 4u \, \sin^2\theta\right]^{\alpha\epsilon-1} (1 + u + 2\sqrt{u}\cos\theta) \tag{A.109}$$

The θ integral can be performed in terms of hypergeometric functions using

$$\int_0^\pi d\theta \, \sin^{-2\epsilon}\theta \left[(1-u)^2 + 4u \, \sin^2\theta\right]^{-1+\alpha\epsilon} = \frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]} \tag{A.110}$$

$$\times (1+u)^{-1+2\epsilon} \left((1-u)^2\right)^{-1/2-(1-\alpha)\epsilon} \, {}_2F_1\left[1/2-\epsilon, -\epsilon+\alpha\epsilon, 1-\epsilon, \frac{4u}{(1+u)^2}\right] \, ,$$

and

$$\int_0^\pi d\theta \, \sin^{-2\epsilon}\theta \left[(1-u)^2 + 4u \, \sin^2\theta\right]^{\alpha\epsilon-1} \cos\theta = 0 \, , \tag{A.111}$$

by symmetry.

The hypergeometric function has the expansion

$${}_2F_1\left[\frac{1}{2} - \epsilon, -(1-\alpha)\epsilon; 1-\epsilon; \frac{4u}{(1+u)^2}\right] = 1 - 2(1-\alpha)\epsilon \ln(1+u) + \mathcal{O}(\epsilon^2) \, , \tag{A.112}$$

which is valid for $0 \leq u \leq 1$,

The final u integral is then

$$\frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]} \int_0^1 du \left[u^{-\epsilon} + u^{-1+(1-2\alpha)\epsilon}\right] (1+u)^{2\epsilon} \left((1-u)^2\right)^{-1/2-(1-\alpha)\epsilon}$$

$$\times \left(1 - 2(1-\alpha)\epsilon \ln(1+u)\right). \qquad \text{(A.113)}$$

We expect this integral to contribute both $\frac{1}{(1-\alpha)\epsilon}$ and $\frac{1}{(1-2\alpha)\epsilon}$ poles, unlike the $n_a n_b$ contribution, which are evident in the $u \to 1$ and $u \to 0$ limits respectively. We need to do the integral to $\mathcal{O}(\epsilon)$ to get the finite pieces, but only $\mathcal{O}(\epsilon^0)$ to get the anomalous dimensions, which is sufficient for now. We have

$$= \frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]} \int_0^1 du\, u^{-\epsilon}(1+u)^{2\epsilon} \left((1-u)^2\right)^{-1/2-(1-\alpha)\epsilon}$$

$$- 2(1-\alpha)\epsilon \frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]} \int_0^1 du\, u^{-\epsilon}(1+u)^{2\epsilon} \left((1-u)^2\right)^{-1/2-(1-\alpha)\epsilon} \log(1+u)$$

$$+ \frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]} \int_0^1 du\, u^{-1+(1-2\alpha)\epsilon}(1+u)^{2\epsilon} \left((1-u)^2\right)^{-1/2-(1-\alpha)\epsilon}$$

$$- 2(1-\alpha)\epsilon \frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]} \int_0^1 du\, u^{-1+(1-2\alpha)\epsilon}(1+u)^{2\epsilon} \left((1-u)^2\right)^{-1/2-(1-\alpha)\epsilon} \log(1+u)$$

$$\text{(A.114)}$$

This integral can be done systematically using +-functions, but to the order we need the result, it is easier to use subtractions, evaluate the log at the value of the singularity, and then perform the integral in terms of hypergeometric functions. The integral can be written

$$= \frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]} \int_0^1 du\, u^{-\epsilon}(1+u)^{2\epsilon} \left((1-u)^2\right)^{-1/2-(1-\alpha)\epsilon} \qquad \text{(A.115)}$$

$$- 2(1-\alpha)\epsilon \frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]} \int_0^1 du\, u^{-\epsilon}(1+u)^{2\epsilon} \left((1-u)^2\right)^{-1/2-(1-\alpha)\epsilon} \log(2)$$

$$+ \frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]} \int_0^1 du\, u^{-1+(1-2\alpha)\epsilon}(1+u)^{2\epsilon} \left((1-u)^2\right)^{-1/2-(1-\alpha)\epsilon}$$

$$- 2(1-\alpha)\epsilon \frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]} \int_0^1 du\, u^{-1+(1-2\alpha)\epsilon}(1+u)^{2\epsilon} \left[\left((1-u)^2\right)^{-1/2-(1-\alpha)\epsilon} - 1\right] \log(2),$$

which gives

$$
= \frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]} \frac{\Gamma[1-\epsilon]\Gamma[-2(1-\alpha)\epsilon]}{\Gamma[1-\epsilon-2(1-\alpha)\epsilon]} \, {}_2F_1[-2\epsilon, 1-\epsilon; 1-\epsilon-2(1-\alpha)\epsilon; -1]
$$

$$
- 2(1-\alpha)\epsilon \frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]} \log(2) \frac{\Gamma[1-\epsilon]\Gamma[-2(1-\alpha)\epsilon]}{\Gamma[1-\epsilon-2(1-\alpha)\epsilon]} \, {}_2F_1[-2\epsilon, 1-\epsilon; 1-\epsilon-2(1-\alpha)\epsilon; -1]
$$

$$
+ \frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]} \frac{\Gamma[(1-2\alpha)\epsilon]\Gamma[-2(1-\alpha)\epsilon]}{\Gamma[(1-2\alpha)\epsilon-2(1-\alpha)\epsilon]} \, {}_2F_1[-2\epsilon, (1-2\alpha)\epsilon; (1-2\alpha)\epsilon-2(1-\alpha)\epsilon; -1]
$$

$$
- 2(1-\alpha)\epsilon \frac{\Gamma[1/2-\epsilon]\Gamma[1/2]}{\Gamma[1-\epsilon]} \log(2)
$$

$$
\times \left(\frac{\Gamma[(1-2\alpha)\epsilon]\Gamma[-2(1-\alpha)\epsilon]}{\Gamma[(1-2\alpha)\epsilon-2(1-\alpha)\epsilon]} \, {}_2F_1[-2\epsilon, (1-2\alpha)\epsilon; (1-2\alpha)\epsilon-2(1-\alpha)\epsilon; -1] \right.
$$

$$
\left. - \frac{\Gamma[(1-2\alpha)\epsilon]}{\Gamma[1+(1-2\alpha)\epsilon]} \, {}_2F_1[-2\epsilon, 1; 1+(1-2\alpha)\epsilon; -1] \right) .
$$

$$(A.116)$$

Expanding this to $\mathcal{O}(\epsilon^0)$ gives

$$
= \frac{\pi}{(\alpha-1)\epsilon} - \frac{\pi}{(2\alpha-1)\epsilon} - \frac{2\pi\log(2)}{2\alpha-1} + \frac{4\pi\log(2)}{\alpha-1} + 2\pi\log(2) \tag{A.117}
$$

We then have

$$
S^{(1)}_{c,n_a\bar{n}}(\tilde{e}^{(\alpha)}_3) = -g^2 4c_\epsilon \Gamma(-2\epsilon) \left(\frac{\mu^2 N^2_{CS} e^{\gamma_E}(\tilde{e}^{(\alpha)}_3)^2 \left(\frac{n_a\cdot n_b}{8}\right)^{-1+2\alpha}}{4\pi 4^{1-\alpha} Q^2} \right)^\epsilon
$$

$$
\times \left(\frac{\pi}{(\alpha-1)\epsilon} - \frac{\pi}{(2\alpha-1)\epsilon} - \frac{2\pi\log(2)}{2\alpha-1} + \frac{4\pi\log(2)}{\alpha-1} + 2\pi\log(2) \right) .
$$

$$(A.118)$$

Extracting just the divergent pieces so as to get the anomalous dimensions, we find

$$
S^{(1)}_{c,n_a\bar{n}}(\tilde{e}^{(\alpha)}_3) = \frac{\alpha_s}{2\pi(\alpha-1)\epsilon^2} - \frac{\alpha_s}{2\pi(2\alpha-1)\epsilon^2} + \frac{\log\left[\frac{\mu N_{CS} e^{\gamma_E}(\tilde{e}^{(\alpha)}_3)\left(\frac{n_a\cdot n_b}{8}\right)^{-1/2+\alpha}}{2^{1-\alpha}Q} \right] \alpha_s}{\pi(\alpha-1)\epsilon}
$$

$$
- \frac{\log\left[\frac{\mu N_{CS} e^{\gamma_E}(\tilde{e}^{(\alpha)}_3)\left(\frac{n_a\cdot n_b}{8}\right)^{-1/2+\alpha}}{2^{1-\alpha}Q} \right] \alpha_s}{\pi(2\alpha-1)\epsilon} + \frac{\log(2)\alpha_s}{\pi(\alpha-1)\epsilon} + \frac{\log(2)\alpha_s}{\pi\epsilon} , \tag{A.119}
$$

387

which can be simplified to

$$S^{(1)}_{c,\,n_a\bar{n}}(\tilde{e}_3^{(\alpha)}) = \frac{\alpha_s}{2\pi(\alpha-1)\epsilon^2} - \frac{\alpha_s}{2\pi(2\alpha-1)\epsilon^2}$$
$$+ \frac{\alpha_s}{\pi(\alpha-1)\epsilon}L^{cs}_\alpha\left(\tilde{e}_3^{(\alpha)}\right) - \frac{\alpha_s}{\pi(2\alpha-1)\epsilon}L^{cs}_\alpha\left(\tilde{e}_3^{(\alpha)}\right), \qquad (A.120)$$

where, as for the logarithm in the $n_a\,n_b$ contribution, Eq. (A.97), the logarithm that appears is

$$L^{cs}_\alpha\left(\tilde{e}_3^{(\alpha)}\right) = \log\left(\frac{\mu N_{CS}e^{\gamma_E}(\tilde{e}_3^{(\alpha)})(n_a\cdot n_b)^{-1/2+\alpha}}{\sqrt{2}Q}\right). \qquad (A.121)$$

The contribution from an emission between the n_b and $\bar{n}$ Wilson lines is identical, so we have

$$S^{(1)}_{c,\,n_b\bar{n}}\left(e_3^{(\alpha)}\right) = S^{(1)}_{c,\,n_a\bar{n}}\left(e_3^{(\alpha)}\right). \qquad (A.122)$$

Note that for both the $\bar{n}\,n_a$ and $\bar{n}\,n_b$ contributions, and unlike for the $n_a\,n_b$ contribution, we have $1/\epsilon$ contributions both of the soft form $1/(1-2\alpha)$, and of the collinear form, $1/(1-\alpha)$. This will be crucial to achieve the cancellation of anomalous dimensions, as required for the consistency of the collinear subjets factorization theorem.

It is interesting to note that this structure is very different than that which appeared for the case of the N-subjettiness observable in Ref. [105]. In this case only a single angular exponent appears throughout the calculation, unlike both the $1/(1-2\alpha)$ and $1/(1-\alpha)$ that we find here, and the divergent pieces of the $\bar{n}\,n_a$ and $\bar{n}\,n_b$ contributions vanish.

Cancellation of Anomalous Dimensions

We now review the renormalization group evolution of each of the functions in the factorization theorem, and show that sum of the anomalous dimensions vanishes, as required for renormalization group consistency.

388

The hard function satisfies a multiplicative RGE, given by

$$\mu\frac{d}{d\mu}\log H(Q^2,\mu) = \gamma_H(Q^2,\mu) = 2\text{Re}\left[\gamma_C(Q^2,\mu)\right] , \tag{A.123}$$

where

$$\gamma_C(Q^2,\mu) = \frac{\alpha_s C_F}{4\pi}\left(4\log\left[\frac{-Q^2}{\mu^2}\right] - 6\right) , \tag{A.124}$$

is the anomalous dimension of the dijet Wilson coefficient. Explicitly

$$\gamma_H(Q^2,\mu) = \frac{\alpha_s C_F}{2\pi}\left(4\log\left[\frac{Q^2}{\mu^2}\right] - 6\right) . \tag{A.125}$$

The anomalous dimension of the hard splitting function H_2 can be extracted from Ref. [105] by performing a change of variables. It satisfies a multiplicative RGE

$$\mu\frac{d}{d\mu}H_2(t,x,\mu) = \gamma_{H_2}(t,x,\mu)H_2(t,x,\mu) , \tag{A.126}$$

with anomalous dimension

$$\gamma_{H_2}(t,x,\mu) = \frac{\alpha_s(\mu)}{2\pi}\left[2C_A\log\frac{t}{\mu^2} + 4\left(C_F - \frac{C_A}{2}\right)\log x + 2C_A\log(1-x) - \beta_0\right] . \tag{A.127}$$

Here β_0 is defined with the normalization

$$\beta_0 = \frac{11C_A}{3} - \frac{2n_f}{3} . \tag{A.128}$$

Converting to $e_2^{(\alpha)}$ by performing the change of variables given in Eq. (A.43), we find

$$\gamma_{H_2}\left(e_2^{(\alpha)},z_q,\mu\right) = \frac{\alpha_s(\mu)}{2\pi}\left[2C_A\log\left(\frac{Q^2}{\mu^2}\frac{(z_a z_b)^{1-2/\alpha}\left(e_2^{(\alpha)}\right)^{2/\alpha}}{4}\right)\right.$$

$$\left. +4\left(C_F - \frac{C_A}{2}\right)\log z_q + 2C_A\log(1-z_q) - \beta_0\right] . \tag{A.129}$$

389

Since the anomalous dimensions of the jet, soft and collinear-soft functions are written in terms of $\tilde{e}_3^{(\alpha)}$, z_a, z_b, and $n_a \cdot n_b$, for demonstrating cancellation of anomalous dimensions, it is convenient to replace $e_2^{(\alpha)}$ in Eq. (A.129) with its leading power expression from Eq. (A.21). We then have

$$\gamma_{H_2}\left(e_2^{(\alpha)}, z_q, \mu\right) = \frac{\alpha_s(\mu)}{2\pi}\left[2C_A\log\left(\frac{Q^2}{\mu^2}\frac{z_a z_b\, n_a \cdot n_b}{2}\right)\right.$$
$$\left. +4\left(C_F - \frac{C_A}{2}\right)\log z_q + 2C_A\log(1 - z_q) - \beta_0\right]. \tag{A.130}$$

Note that $1 - z_q = z_g$.

The jet functions satisfy multiplicative RGEs in Laplace space (they satisfy convolutional RGEs in $e_3^{(\alpha)}$, see Ref. [167] for a detailed discussion)

$$\mu\frac{d}{d\mu}\log \tilde{J}_{g,qn}\left(Q_J, \tilde{e}_3^{(\alpha)}\right) = \gamma_{g,q}^\alpha\left(Q_J, \tilde{e}_3^{(\alpha)}\right), \tag{A.131}$$

where the one-loop anomalous dimension is determined from Eqs. (A.68) and (A.69), and is given by

$$\gamma_{g,q}^\alpha\left(Q_J, \tilde{e}_3^{(\alpha)}\right) = -2\frac{\alpha_s}{\pi}\frac{C_{g,q}}{(1-\alpha)}L_\alpha^{J,a}\left(\tilde{e}_3^{(\alpha)}\right) + \gamma_{g,q}, \tag{A.132}$$

where the logarithm $L_\alpha^{J,a}\left(\tilde{e}_3^{(\alpha)}\right)$ was defined in Eq. (A.70), and is given by

$$L_\alpha^{J,a}\left(\tilde{e}_3^{(\alpha)}\right) = \log\left[N_J\tilde{e}_3^{(\alpha)}e^{\gamma_E}\left(\frac{\mu}{\sqrt{2}Q}\right)^\alpha z_a^{2-\alpha}\right]. \tag{A.133}$$

Here $C_{g,q}$ is the appropriate Casimir (C_A for gluon jets and C_F for quark jets), and with $\gamma_{g,q}$ the standard functions

$$\gamma_q = \frac{3\alpha_s C_F}{2\pi}, \qquad \gamma_g = \frac{\alpha_s}{\pi}\frac{11C_A - 2n_f}{6}. \tag{A.134}$$

For subjet b, we simply have $a \to b$.

390

Similarly, the soft function satisfies a multiplicative RGE in Laplace space

$$\mu \frac{d}{d\mu} \log \tilde{S}_G \left(\tilde{e}_3^{(\alpha)} \right) = \gamma_G \left(\tilde{e}_3^{(\alpha)} \right) , \qquad (A.135)$$

with one-loop anomalous dimension determined by Eq. (A.58), and given by

$$\gamma_G \left(\tilde{e}_3^{(\alpha)} \right) = \frac{-2\alpha_s}{\pi(1-2\alpha)} \mathbf{T}_n \cdot \mathbf{T}_{\bar{n}} \, L_\alpha^G \left(\tilde{e}_3^{(\alpha)} \right) . \qquad (A.136)$$

Here the logarithm is given by

$$L_\alpha^G \left(\tilde{e}_3^{(\alpha)} \right) = \log \left[\frac{e^{\gamma_E} \mu \tilde{e}_3^{(\alpha)} N_S}{2^{1-\alpha} Q} \right] - \frac{(1-2\alpha)}{2} \log \left[\tan^2 \frac{R}{2} \right] . \qquad (A.137)$$

Finally, the collinear soft function satisfies a multiplicative RGE in Laplace space

$$\mu \frac{d}{d\mu} \log S_c \left(\tilde{e}_3^{(\alpha)} \right) = \gamma_{cs} \left(\tilde{e}_3^{(\alpha)} \right) , \qquad (A.138)$$

with the one-loop anomalous dimension determined by Eqs. (A.96) and (A.119)

$$\gamma_{cs} \left(\tilde{e}_3^{(\alpha)} \right) = \mathbf{T}_a \cdot \mathbf{T}_b \gamma_{ab} \left(\tilde{e}_3^{(\alpha)} \right) + \mathbf{T}_a \cdot \mathbf{T}_{\bar{n}} \gamma_{a\bar{n}} \left(\tilde{e}_3^{(\alpha)} \right) + \mathbf{T}_{\bar{n}} \cdot \mathbf{T}_b \gamma_{\bar{n}b} \left(\tilde{e}_3^{(\alpha)} \right) , \qquad (A.139)$$

where

$$\gamma_{ab} \left(\tilde{e}_3^{(\alpha)} \right) = \frac{-4\alpha_s}{\pi(1-\alpha)} L_\alpha^{cs} \left(\tilde{e}_3^{(\alpha)} \right) , \qquad (A.140)$$

$$\gamma_{a\bar{n}} \left(\tilde{e}_3^{(\alpha)} \right) = \gamma_{b\bar{n}} \left(\tilde{e}_3^{(\alpha)} \right) = \frac{-2\alpha_s}{\pi(1-\alpha)} L_\alpha^{cs} \left(\tilde{e}_3^{(\alpha)} \right) + \frac{2\alpha_s}{\pi(1-2\alpha)} L_\alpha^{cs} \left(\tilde{e}_3^{(\alpha)} \right) . \qquad (A.141)$$

The argument of the logarithm appearing in the collinear soft function, was defined in Eq. (A.97), and is given by

$$L_\alpha^{cs} = \log \left(\frac{\mu N_{CS} e^{\gamma_E} (\tilde{e}_3^{(\alpha)}) (n_a \cdot n_b)^{-1/2+\alpha}}{\sqrt{2} Q} \right) . \qquad (A.142)$$

391

We can now explicitly check the cancellation of anomalous dimensions. We consider the particular partonic subprocess $e^+e^- \to \bar{q}q \to \bar{q}qg$ for which we have explicitly given the hard splitting function, in which case the color algebra can be simplified and written entirely in terms of Casimirs using the color conservation relations

$$\mathbf{T}_n = \mathbf{T}_q + \mathbf{T}_g \,, \tag{A.143}$$

$$\mathbf{T}_n + \mathbf{T}_{\bar{n}} = 0 \,. \tag{A.144}$$

We then have

$$\mathbf{T}_n \cdot \mathbf{T}_{\bar{n}} = -C_F \,, \tag{A.145}$$

$$\mathbf{T}_q \cdot \mathbf{T}_g = -\frac{C_A}{2} \,, \tag{A.146}$$

$$\mathbf{T}_q \cdot \mathbf{T}_{\bar{n}} = \frac{C_A}{2} - C_F \,, \tag{A.147}$$

$$\mathbf{T}_g \cdot \mathbf{T}_{\bar{n}} = -\frac{C_A}{2} \,, \tag{A.148}$$

$$\mathbf{T}_g \cdot \mathbf{T}_g = C_A \,, \tag{A.149}$$

$$\mathbf{T}_n \cdot \mathbf{T}_n = C_F \,. \tag{A.150}$$

However, for most of the cancellation of the anomalous dimensions, it will be convenient to work in the abstract color notation, so as not to need to use relations between the color Casimirs.

The independence of the total cross section under renormalization group evolution implies the following relation between anomalous dimensions

$$\gamma_H \left(Q^2, \mu\right) + \gamma_{H_2} \left(e_2^{(\alpha)}, z_q, \mu\right) + \gamma_g^\alpha \left(\tilde{e}_3^{(\alpha)}\right) + \gamma_q^\alpha \left(\tilde{e}_3^{(\alpha)}\right) + \gamma_G \left(\tilde{e}_3^{(\alpha)}\right) + \gamma_{cs} \left(\tilde{e}_3^{(\alpha)}\right) \sim 0 \,, \tag{A.151}$$

where the $\sim$ means up to a term corresponding to the measurement of the jet in the $\bar{n}$ direction, and the out-of-jet contribution to the soft function, which is independent of the

$\tilde{e}_3^{(\alpha)}$ measurement, and the kinematics of the substructure, namely $n_a \cdot n_b$, z_a, and z_b. We will make this relation precise shortly.

We now show explicitly that this cancellation occurs, and how it arises, which provides a non-trivial cross-check on the collinear-subjets factorization theorem. Substituting in the expressions above, we find

$$
\sum_\gamma = \gamma_H \left(Q^2, \mu \right) + \gamma_{H_2} \left(e_2^{(\alpha)}, z_q, \mu \right)
$$
$$
+ \left[-\mathbf{T}_a \cdot \mathbf{T}_b \frac{4\alpha_s L_\alpha^{cs} \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-\alpha)} \right.
$$
$$
\left. -2\mathbf{T}_a \cdot \mathbf{T}_{\bar{n}} \left(\frac{\alpha_s L_\alpha^{cs} \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-\alpha)} - \frac{\alpha_s L_\alpha^{cs} \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-2\alpha)} \right) - 2\mathbf{T}_b \cdot \mathbf{T}_{\bar{n}} \left(\frac{\alpha_s L_\alpha^{cs} \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-\alpha)} - \frac{\alpha_s L_\alpha^{cs} \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-2\alpha)} \right) \right]
$$
$$
- \left[\mathbf{T}_n \cdot \mathbf{T}_{\bar{n}} \frac{2\alpha_s L_\alpha^G \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-2\alpha)} \right] - C_A \frac{2\alpha_s L_\alpha^g \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-\alpha)} + \gamma_g - C_F \frac{2\alpha_s L_\alpha^q \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-\alpha)} + \gamma_q . \tag{A.152}
$$

To make manifest the separate cancellations, we use the color conservation relation $\mathbf{T}_n = \mathbf{T}_a + \mathbf{T}_b$ in the soft anomalous dimension, and $\mathbf{T}_{\bar{n}} = -\mathbf{T}_a - \mathbf{T}_b$ in the $1/(1-\alpha)$ pieces of the collinear soft anomalous dimensions. Grouping together collinear like terms $(1/(1-\alpha))$ and soft like terms $(1/(1-2\alpha))$, we then have

$$
\sum_\gamma = \gamma_H \left(Q^2, \mu \right) + \gamma_{H_2} \left(e_2^{(\alpha)}, z_q, \mu \right)
$$
$$
- \left[(\mathbf{T}_a + \mathbf{T}_b) \cdot \mathbf{T}_{\bar{n}} \frac{2\alpha_s L_\alpha^G \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-2\alpha)} \right] + \left[\mathbf{T}_a \cdot \mathbf{T}_{\bar{n}} \frac{2\alpha_s L_\alpha^{cs} \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-2\alpha)} + \mathbf{T}_b \cdot \mathbf{T}_{\bar{n}} \frac{2\alpha_s L_\alpha^{cs} \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-2\alpha)} \right]
$$
$$
- \mathbf{T}_a \cdot \mathbf{T}_b \frac{4\alpha_s L_\alpha^{cs} \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-\alpha)} - \left[\mathbf{T}_a \cdot (-\mathbf{T}_a - \mathbf{T}_b) \frac{2\alpha_s L_\alpha^{cs} \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-\alpha)} + \mathbf{T}_b \cdot (-\mathbf{T}_a - \mathbf{T}_b) \frac{2\alpha_s L_\alpha^{cs} \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-\alpha)} \right]
$$
$$
- C_A \frac{2\alpha_s L_\alpha^g \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-\alpha)} + \gamma_g - C_F \frac{2\alpha_s L_\alpha^q \left(\tilde{e}_3^{(\alpha)} \right)}{\pi(1-\alpha)} + \gamma_q . \tag{A.153}
$$

393

Since all the logs are linear in the $\tilde{e}_3^{(\alpha)}$, we immediately see that the color conservation relations have led to the cancellation of the $\tilde{e}_3^{(\alpha)}$ dependence in the soft like pieces between the $\bar{n}n_b$ and $\bar{n}n_a$ contributions to the collinear soft function with the global soft contribution, and the cancellation between the collinear like pieces involve all three contributions to the collinear soft function, as well as the jet functions. This nontrivial cancellation supports the validity of the collinear subjets factorization theorem.

It is also straightforward to check that the dependence on $e_2^{(\alpha)}$ as well as on the jet energy fractions also cancels, although this is more tedious to perform step by step. We therefore simply quote the summed result of the anomalous dimensions, to make clear the meaning of the equivalence relation in Eq. (A.151). We have

$$\gamma_H\left(Q^2,\mu\right) + \gamma_{H_2}\left(e_2^{(\alpha)}, z_q, \mu\right) + \gamma_g^\alpha\left(\tilde{e}_3^{(\alpha)}\right) + \gamma_q^\alpha\left(\tilde{e}_3^{(\alpha)}\right) + \gamma_G\left(\tilde{e}_3^{(\alpha)}\right) + \gamma_{cs}\left(\tilde{e}_3^{(\alpha)}\right) =$$
$$-\frac{3\alpha_s C_F}{2\pi} - \frac{\alpha_s C_F \log\left[\tan^2\frac{R}{2}\right]}{\pi} - \frac{\alpha_s C_F \log\frac{\mu^2}{Q^2}}{\pi}. \tag{A.154}$$

These remaining terms are exactly those expected to cancel against the out-of-jet contribution; see, e.g., Ref. [167] for a detailed discussion.

The out-of-jet jet function is then given by the unmeasured jet function of Ref. [167]

$$\mu\frac{d}{d\mu}\ln J_{oj}(R_B) = \frac{2\alpha_s C_F}{\pi}\log\left[\frac{\mu}{Q\tan\frac{R_B}{2}}\right] + \frac{3\alpha_s C_F}{2\pi}, \tag{A.155}$$

where here R_B is the radius of the recoiling jet. For simplicity, throughout this chapter, we have taken $R_B = R$.

The out-of-jet contribution to the soft function has a pure cusp anomalous dimension [167]

$$\mu\frac{d}{d\mu}\ln S_{oj}(R_B) = \frac{2\alpha_s C_F}{\pi}\log\left[\tan^2\frac{R}{2}\right]. \tag{A.156}$$

394

A.3 One Loop Calculations of Soft Subjet Functions

In this appendix we give the operator definitions and one-loop results for the functions appearing in the factorization theorem of Eq. (4.27) for the soft subjet region of phase space. The factorization theorem in the soft subjet region of phase space was first presented in Ref. [11], where all functions were calculated to one-loop, and a detailed discussion of the structure of the required zero bin subtractions was given. This calculation was performed with a broadening axis cone algorithm, however it was argued in Sec. 4.3.1 that to leading power, the factorization theorem is identical in the case of an anti-k_T algorithm. Because of this, in this appendix we give only the final results for the one-loop anomalous dimensions, and the tree level matching for the soft subjet production, as are required for the resummation considered in this chapter. The interested reader is referred to Ref. [11] for the detailed calculation, as well as a discussion of the intricate zero bin structure of the factorization theorem, which is only briefly mentioned in this appendix.

Definitions of Factorized Functions

The functions appearing in the soft subjet factorization theorem of Eq. (4.27) have the following SCET operator definitions:

- Hard Matching Coefficient for Dijet Production

$$H(Q^2, \mu) = |C(Q^2, \mu)|^2 , \qquad (A.157)$$

where $C(Q^2, \mu)$ is the Wilson coefficient obtained from matching the full theory QCD current $\bar{\psi}\Gamma\psi$ onto the SCET dijet operator $\bar{\chi}_n\Gamma\chi_{\bar{n}}$

$$\langle q\bar{q}|\bar{\psi}\Gamma\psi|0\rangle = C(Q^2, \mu) \langle q\bar{q}|\mathcal{O}_2|0\rangle . \qquad (A.158)$$

As before, we have neglected the contraction with the Leptonic tensor.

- **Soft Subject Jet Function:**

$$J_{n_{sj}}\left(e_3^{(\alpha)}\right) = \tag{A.159}$$

$$\frac{(2\pi)^3}{C_A}\mathrm{tr}\langle 0|\mathcal{B}^\mu_{\perp_{n_{sj}}}(0)\Theta_O(B)\delta(Q_{SJ}-\bar{n}_{sj}\cdot\mathcal{P})\delta^{(2)}(\vec{\mathcal{P}}_{\perp_{SJ}})\delta\left(e_3^{(\alpha)}-\Theta_{FJ}\mathbf{E_3}^{(\alpha)}\big|_{SJ}\right)\mathcal{B}_{\perp_{n_{sj}}\mu}(0)|0\rangle$$

- **Jet Function:**

$$J_n\left(e_3^{(\alpha)}\right) = \frac{(2\pi)^3}{C_F}\mathrm{tr}\langle 0|\frac{\slashed{\bar{n}}}{2}\chi_n(0)\Theta_O(B)\delta(Q-\bar{n}\cdot\mathcal{P})\delta^{(2)}(\vec{\mathcal{P}}_\perp)\delta\left(e_3^{(\alpha)}-\Theta_{FJ}\mathbf{E_3}^{(\alpha)}\big|_{HJ}\right)\bar{\chi}_n(0)|0\rangle$$

$$\tag{A.160}$$

- **Boundary Soft Function:**

$$S_{n_{sj}\bar{n}_{sj}}\left(e_3^{(\alpha)};R\right) = \frac{1}{C_A}\mathrm{tr}\langle 0|T\{S_{n_{sj}}S_{\bar{n}_{sj}}\}\Theta_O(B)\delta\left(e_3^{(\alpha)}-\Theta_{FJ}\mathbf{E_3}^{(\alpha)}\big|_{BS}\right)\bar{T}\{S_{n_{sj}}S_{\bar{n}_{sj}}\}|0\rangle$$

$$\tag{A.161}$$

- **Soft Subject Soft Function:**

$$S_{n_{sj}n\bar{n}}\left(e_3^{(\alpha)},B;R\right) = \mathrm{tr}\langle 0|T\{S_{n_{sj}}S_n S_{\bar{n}}\}\Theta_O(B)\delta\left(e_3^{(\alpha)}-\Theta_{FJ}\mathbf{E_3}^{(\alpha)}\big|_S\right)\bar{T}\{S_{n_{sj}}S_n S_{\bar{n}}\}|0\rangle$$

$$\tag{A.162}$$

The definitions of these functions include measurement operators, which when acting on the final state, return the value of a given observable. The operator $\mathbf{E_3}^{(\alpha)}$ measures the contribution to $e_3^{(\alpha)}$ from final states, and must be appropriately expanded following the power counting of the sector on which it acts. Expressions for the expansions in the power counting of the different sectors will be given shortly, after kinematic notation has been set up. The operators Θ_{FJ}, and Θ_O constrain the measured radiation to be in the jet or out of the jet, respectively, and will be defined shortly.

Kinematics and Notation

For our general kinematic setup, we will denote by Q the center of mass energy of the e^+e^- collisions, so that $Q/2$ is the energy deposited in a hemisphere. i.e. the four-momenta of the two hemispheres are

$$p_{\text{hemisphere}_1} = \left(\frac{Q}{2}, \vec{p}_1\right), \qquad p_{\text{hemisphere}_2} = \left(\frac{Q}{2}, -\vec{p}_1\right) \qquad (A.163)$$

so

$$s = Q^2. \qquad (A.164)$$

We are now interested in the regime where there is a wide angle soft subjet carrying a small energy fraction, and an energetic subjet, carrying the majority of the energy fraction. We will label the lightcone directions of the energetic subjet by $n, \bar{n}$, and the lightcone directions of the soft subjet as $n_{sj}, \bar{n}_{sj}$. We will use the variable z_{sj} to label the energy fraction of the soft subjet, namely

$$E_{sj} = z_{sj}\frac{Q}{2}, \qquad z_{sj} \ll 1. \qquad (A.165)$$

In this region of phase space, to leading power the value of the two point energy correlation function is set by the two subjets, and is given by

$$e_2^{(\alpha)} = 2^{\alpha/2} z_{sj} \left(n \cdot n_{sj}\right)^{\alpha/2}. \qquad (A.166)$$

The action of the measurement function $\mathbf{E_3}^{(\alpha)}$ on a arbitrary state for each of the factorized sectors contributing to the 3-point energy correlation function measurement is given by

$$\mathbf{E_3}^{(\alpha)}\big|_{SJ}\big|X_{sj}\big\rangle = \sum_{k_i, k_j \in X_{sj}} N_{SJ}\frac{\bar{n}_{sj} \cdot k_i}{Q}\frac{\bar{n}_{sj} \cdot k_j}{Q}\left(\frac{k_i \cdot k_j}{\bar{n}_{sj} \cdot k_i \bar{n}_{sj} \cdot k_j}\right)^{\frac{\alpha}{2}}\big|X_{sj}\big\rangle, \qquad (A.167)$$

397

$$\mathbf{E_3}^{(\alpha)}\Big|_{HJ}\Big|X_{hj}\Big\rangle = \sum_{k_i,k_j \in X_{hj}} N_{HJ} \frac{\bar{n} \cdot k_i}{Q} \frac{\bar{n} \cdot k_j}{Q} \left(\frac{k_i \cdot k_j}{\bar{n} \cdot k_i \bar{n} \cdot k_j}\right)^{\frac{\alpha}{2}} \Big|X_{hj}\Big\rangle, \qquad \text{(A.168)}$$

$$\mathbf{E_3}^{(\alpha)}\Big|_{BS}\Big|X_{bs}\Big\rangle = \sum_{k \in X_{bs}} N_{BS} \frac{\bar{n}_{sj} \cdot k}{Q} \left(\frac{n_{sj} \cdot k}{\bar{n}_{sj} \cdot k}\right)^{\frac{\alpha}{2}} \Big|X_{bs}\Big\rangle, \qquad \text{(A.169)}$$

$$\mathbf{E_3}^{(\alpha)}\Big|_{S}\Big|X_{s}\Big\rangle = \sum_{k \in X_{s}} N_{S} \frac{k^0}{Q} \left(\frac{n_{sj} \cdot k}{k^0} \frac{n \cdot k}{k^0}\right)^{\frac{\alpha}{2}} \Big|X_{s}\Big\rangle, \qquad \text{(A.170)}$$

where, for simplicity, we have extracted the normalization factors

$$N_{SJ} = 2^{5\alpha/2}(n \cdot n_{sj})^{\alpha}, \qquad N_{HJ} = 2^{5\alpha/2} z_{sj}(n \cdot n_{sj})^{\alpha}, \qquad \text{(A.171)}$$

$$N_{BS} = 2^{2\alpha} z_{sj}(n \cdot n_{sj})^{\alpha}, \qquad N_{S} = 2^{1+3\alpha/2} z_{sj}(n \cdot n_{sj})^{\alpha/2}. \qquad \text{(A.172)}$$

These expressions follow from properly expanding the definition of the energy correlation function measurements in the power counting of each of the sectors. Note that on the jet sectors, the 3-point correlation measurement becomes an effective 2-point correlation measurement, since the 2-point energy correlation function is set by the initial splitting of the subjet.

The in-jet restriction, Θ_{FJ}, is given by

$$\Theta_{FJ}(k) = \Theta\left(\tan^2\frac{R}{2} - \frac{n \cdot k}{\bar{n} \cdot k}\right). \qquad \text{(A.173)}$$

The jet restriction must also be expanded following the power counting of the given sector. We will see that this is actually quite subtle for the soft subjet modes, since the angle between the soft subjet axis and the boundary of the jet has a non-trivial power counting. In particular, the expansion of $\Theta_{FJ}(k)$ is different for the soft subjet jet and boundary soft modes, and will demonstrate the necessity of performing the complete factorization of the soft subjet dynamics into jet and boundary soft modes. Finally, since we are considering the case where the out-of-jet scale B is much less than the in-jet scale, the operator

$$\Theta_O(B)$$

398

must also be included in the definition of the soft subjet functions. This operators vetoes out-of-jet radiation above the scale B. The explicit expression for $\Theta_O(B)$ expanded in the power counting of each of the factorized sectors can be found in Ref. [11].

Hard Matching Coefficient for Dijet Production

The hard matching coefficient for dijet production, $H(Q^2, \mu)$, is identical to that for the collinear subjets factorization theorem by hard-collinear-soft factorization, and is given in Eq. (A.38).

Hard Matching for Soft Subjet Production

The hard matching coefficient $H^{sj}(z_{sj}, \theta_{sj})$ is determined by the finite parts of the logarithm of the soft matrix element for a single soft state

$$H^{sj}(z_{sj}, n_{sj}) = \text{tr}\langle 0|T\{S_n S_{\bar{n}}\}|sj\rangle\langle sj|\bar{T}\{S_n S_{\bar{n}}\}|0\rangle_{\text{fin}}. \tag{A.174}$$

The virtual corrections of the effective theory cancel the IR divergences of this matrix element, giving a finite matching coefficient. This matrix element can be calculated from the square of the soft gluon current [482, 483], which is known to two loop order [484, 485]. The tree level hard matching coefficient for the soft subjet production is given by

$$H^{sj(\text{tree})}_{n\bar{n}}(z_{sj}, n_{sj}) = \frac{\alpha_s C_F}{\pi z_{sj}} \frac{n \cdot \bar{n}}{n \cdot n_{sj}\, n_{sj} \cdot \bar{n}}. \tag{A.175}$$

The results of Ref. [483] can be used to determine the soft subjet production matching from an arbitrary number of hard jets at one loop.

Anomalous Dimensions

In this section we collect the one-loop anomalous dimensions for all the functions calculated in this appendix. The two hard functions satisfy multiplicative renormalization group

equations. For the dijet production hard function, we have

$$\mu\frac{d}{d\mu}\ln H(Q^2,\mu) = \gamma_H(Q^2,\mu) = 2\mathrm{Re}\left[\gamma_C(Q^2,\mu)\right] . \tag{A.176}$$

Explicitly

$$\gamma_H(Q^2,\mu) = \frac{\alpha_s C_F}{2\pi}\left(4\log\left[\frac{Q^2}{\mu^2}\right] - 6\right) . \tag{A.177}$$

For the soft subjet production hard function, we have

$$\mu\frac{d}{d\mu}\ln H^{sj}_{n\bar{n}}(z_{sj},n_{sj},\mu) = -\frac{\alpha_s C_A}{\pi}\ln\left[\frac{2\mu^2\bar{n}\cdot n}{Q^2 z_{sj}^2 n\cdot n_{sj}\, n_{sj}\cdot\bar{n}}\right] - \frac{\alpha_s}{2\pi}\beta_0 . \tag{A.178}$$

The jet, boundary soft, and global soft functions satisfy multiplicative renormalization group equations in Laplace space, where the Laplace conjugate variable to $e_3^{(\alpha)}$ will be denoted $\tilde{e}_3^{(\alpha)}$

The jet function for the soft subjet satisfies the RGE

$$\mu\frac{d}{d\mu}\ln J_{n_{sj}}\left(\tilde{e}_3^{(\alpha)}\right) = -4\frac{\alpha_s C_A}{2\pi(1-\alpha)}\log\left[2^{-\alpha/2}\tilde{e}_3^{(\alpha)}e^{\gamma_E}z_{sj}^2\left(\frac{\mu}{z_{sj}Q}\right)^\alpha N_{SJ}\right] + \frac{\alpha_s}{2\pi}\beta_0 , \tag{A.179}$$

where the normalization factor N_{SJ} was defined in Eq. (A.171). We have assumed that the soft subjet is a gluon jet, as it is this case that exhibits the soft singularity of QCD.

The jet function for the hard subjet, which we have assumed to be a quark jet, satisfies the RGE

$$\mu\frac{d}{d\mu}\ln J_{hj}\left(\tilde{e}_3^{(\alpha)}\right) = -4\frac{\alpha_s C_F}{2\pi(1-\alpha)}\log\left[2^{-\alpha/2}\tilde{e}_3^{(\alpha)}e^{\gamma_E}\left(\frac{\mu}{Q}\right)^\alpha N_{HJ}\right] + \frac{3\alpha_s C_F}{2\pi} . \tag{A.180}$$

where the normalization factor N_{HJ} was defined in Eq. (A.171).

Since the soft subjet factorization theorem is sensitive to the boundary of the jet, it is also necessary to include out-of-jet contributions. We assume that nothing is measured on the recoiling jet. The out-of-jet jet function is then given by the unmeasured jet function of

400

$$\mu \frac{d}{d\mu} \ln J_{oj}(R_B) = \frac{2\alpha_s C_F}{\pi} \log \left[\frac{\mu}{Q \tan \frac{R_B}{2}} \right] + \frac{3\alpha_s C_F}{2\pi} , \qquad \text{(A.181)}$$

where here R_B is the radius of the recoiling jet. For simplicity, throughout this chapter, we have taken $R_B = R$.

The boundary soft function, satisfies the RGE

$$\mu \frac{d}{d\mu} \ln S_{n_{sj} \bar{n}_{sj}} \left(\tilde{e}_3^{(\alpha)}; R \right) = \frac{\alpha_s C_A}{\pi(1-\alpha)} \log \left[\frac{\mu}{Q} \tilde{e}_3^{(\alpha)} e^{\gamma_E} 2^{1-\alpha} N_{BS} \left(\frac{\bar{n} \cdot n_{sj}}{n \cdot n_{sj}} \tan^4 \frac{R}{2} \right)^{\frac{-(1-\alpha)}{2}} \right.$$
$$\left. \left(1 - \frac{n \cdot n_{sj}}{\bar{n} \cdot n_{sj} \tan^2 \frac{R}{2}} \right)^{-(1-\alpha)} \right] . \qquad \text{(A.182)}$$

where the normalization factor N_{BS} was defined in Eq. (A.171).

For the soft function, it is necessary to perform a refactorization into in-jet and out-of-jet contributions along the lines of Ref. [167]. This is particularly important in the present case, since as was discussed in detail in Ref. [11], the out-of-jet contribution to the soft function is sensitive to the large logarithm, $\log \left[\tan^2 \frac{R}{2} - \tan^2 \frac{\theta_{sj}}{2} \right]$, but due to zero bin subtractions, the in-jet contribution to the soft function does not exhibit such a sensitivity.

The in-jet anomalous dimension has both C_A and C_F contributions. It is given by

$$\gamma_{GS}^{(\text{in})} = - \left(\frac{C_A}{2} - C_F \right) \left(\frac{2\alpha_s}{\pi(1-\alpha)} \log[T] - \frac{2\alpha_s}{\pi} \log \left[\frac{2 \tan \frac{R}{2}}{\tan \frac{\theta_{sj}}{2}} \right] \right)$$
$$- \left(\frac{-C_A}{2} \right) \left(\frac{3\alpha_s}{\pi(1-\alpha)} \log[T] + \frac{2\alpha_s}{\pi} \log \left[\frac{1}{2\bar{n} \cdot n_{sj}} \frac{\tan \frac{R}{2}}{\tan \frac{\theta_{sj}}{2}} \right] \right)$$
$$- \left(\frac{-C_A}{2} \right) \left(\frac{\alpha_s}{\pi(1-\alpha)} \log[T] - \frac{2\alpha_s}{\pi} \log \left[\bar{n} \cdot n_{sj} \right] \right) \qquad \text{(A.183)}$$
$$= (C_A + 2C_F) \frac{\alpha_s}{\pi(1-\alpha)} \log[T] + \frac{C_A \alpha_s}{\pi} \log \left[\frac{\tan^2 \frac{R}{2}}{(\bar{n} \cdot n_{sj})^2 \tan^2 \frac{\theta_{sj}}{2}} \right]$$
$$- \frac{2C_F \alpha_s}{\pi} \log \left[\frac{2 \tan \frac{R}{2}}{\tan \frac{\theta_{sj}}{2}} \right] , \qquad \text{(A.184)}$$

where in the first equality we have separated the contributions from a gluon between the three different Wilson lines, and to simplify the expression we have extracted the argument of the logs

$$T = e^{\gamma_E} N_S \frac{\tilde{e}_3^{(\alpha)} \mu}{Q \tan^{1-\alpha}\frac{\theta_{sj}}{2}} \left(\frac{n \cdot n_{sj}}{2}\right)^{\alpha/2} . \tag{A.185}$$

We choose the canonical scale for the in-jet soft function by minimizing the arguments of the C_A log. Namely, we rewrite the anomalous dimension as

$$\gamma_{GS}^{(\text{in})} = \frac{(C_A + 2C_F)\alpha_s}{\pi(1-\alpha)} \log \left[T \left(\frac{\tan\frac{R}{2}}{(\bar{n} \cdot n_{sj})\tan\frac{\theta_{sj}}{2}} \right)^{2(1-\alpha)} \right]$$

$$+ \frac{2C_F \alpha_s}{\pi(1-\alpha)} \log \left[\left(\frac{\tan\frac{\theta_{sj}}{2}}{2\tan\frac{R}{2}} \right)^{1-\alpha} \left(\frac{\tan\frac{R}{2}}{(\bar{n} \cdot n_{sj})\tan\frac{\theta_{sj}}{2}} \right)^{-2(1-\alpha)} \right] . \tag{A.186}$$

The argument of the second logarithm is formally an $\mathcal{O}(1)$ number in the soft subjet region of phase space, and is treated as the non-cusp anomalous dimension. The argument of the first logarithm is used to set the scale.

The out-of-jet anomalous dimension is purely non-cusp, and is given by

$$\gamma_{GS}^{(\text{out})} = - \left(\frac{C_A}{2} - C_F \right) \frac{2\alpha_s}{\pi} \log \left[\tan\frac{R}{2} \tan\frac{R_B}{2} \right] \tag{A.187}$$

$$- \frac{\alpha_s C_A}{2\pi} \log \left[\frac{\tan^2\frac{R}{2}}{\tan^2\frac{R}{2} - \tan^2\frac{\theta_{sj}}{2}} \right] - \frac{\alpha_s C_A}{2\pi} \log \left[\frac{1}{\tan^2\frac{R_B}{2} \left(\tan^2\frac{R}{2} - \tan^2\frac{\theta_{sj}}{2} \right)} \right] .$$

The natural scale for the out-of-jet soft function is

$$\mu_{\text{out}} = \frac{2n \cdot n_{sj} B}{\tan\frac{\theta_{sj}}{2}} , \tag{A.188}$$

where B is the out-of-jet scale. We set $B = Q \left(e_2^{(\alpha)} \right)^2$ as discussed in Sec. 4.3.1.

For consistency of our soft subjet factorization theorem, the sum of the anomalous di-

402

mensions listed above should cancel. Indeed, one can explicitly check that the anomalous dimensions satisfy the consistency condition

$$
\mu \frac{d}{d\mu} \ln H(Q^2, \mu) + \mu \frac{d}{d\mu} \ln H_{n\bar{n}}^{sj}(z_{sj}, n_{sj}, \mu) + \mu \frac{d}{d\mu} \ln J_{n_{sj}} \left(\tilde{e}_3^{(\alpha)} \right) + \mu \frac{d}{d\mu} \ln J_{hj} \left(\tilde{e}_3^{(\alpha)} \right)
$$
$$
+ \mu \frac{d}{d\mu} \ln J_{oj}(R_B) + \mu \frac{d}{d\mu} \ln S_{n_{sj} \bar{n}_{sj}} \left(\tilde{e}_3^{(\alpha)}; R \right) + \mu \frac{d}{d\mu} \ln S_{n_{sj} n\bar{n}} \left(\tilde{e}_3^{(\alpha)}, B; R, R_B \right) = 0 . \quad \text{(A.189)}
$$

This cancellation is highly non-trivial, involving intricate cancellations between a large number of scales, providing support for the structure of our factorization theorem at the one-loop level. Some further details on the structure of the cancellations, particularly on the dependence of the angle between the soft subjet axis and the boundary, are discussed in Ref. [11].

A.4 Soft Subjet Collinear Zero Bin

In this appendix we summarize the one-loop anomalous dimensions, and required tree level matrix elements for the calculation of the collinear zero bin of the soft subjet factorization theorem, which is required to interpolate between the collinear subjets and soft subjets factorization theorem. Although all the ingredients in this appendix can be obtained straightforwardly from App. A.3 using the standard zero bin procedure [176], we explicitly summarize the results here for completeness.

To perform the zero-bin, all anomalous dimensions and matrix elements of the soft subjet factorization theorem are written in terms of $e_2^{(\alpha)}$ and z_{sj}, and then the limit

$$
\frac{e_2^{(\alpha)}}{z_{sj}} \to 0 \quad \text{(A.190)}
$$

is taken. We will therefore write the anomalous dimensions and matrix elements in this section in terms of $e_2^{(\alpha)}$, z_{sj}, and $\tilde{e}_3^{(\alpha)}$. To keep the notation as simple as possible, we will use only a tilde to denote a collinear zero binned matrix element or anomalous dimension, e.g. $\gamma_{GS}^{(in)} \to \tilde{\gamma}_{GS}^{(in)}$.

Hard Matching for Soft Subjet Production

The collinear binned hard matching coefficient for soft subjet production is given at tree level by

$$\tilde{H}_{n\bar{n}}^{sj(\text{tree})}\left(z_{sj}, e_2^{(\alpha)}\right) = \frac{\alpha_s C_F}{\pi} \frac{2}{\alpha} \frac{1}{z_{sj} e_2^{(\alpha)}} . \tag{A.191}$$

Anomalous Dimensions

Since the renormalization group evolution of all functions in the zero bin is identical to in the soft subjet factorization theorem, here we simply list the results for the zero binned one-loop anomalous dimensions:

$$\gamma_{\tilde{H}} = \frac{\alpha_s C_F}{2\pi}\left(4\log\left[\frac{Q^2}{\mu^2}\right] - 6\right) , \tag{A.192}$$

$$\gamma_{\tilde{H}_{n\bar{n}}^{sj}} = -\frac{2C_A \alpha_s}{\pi}\log\left[\frac{2\mu}{Q z_{sj}^{1-1/\alpha}\left(e_2^{(\alpha)}\right)^{1/\alpha}}\right] - \frac{\alpha_s}{2\pi}\beta_0 , \tag{A.193}$$

$$\gamma_{\tilde{J}_{oj}} = \frac{2\alpha_s C_F}{\pi}\log\left[\frac{\mu}{Q\tan\frac{R_B}{2}}\right] + \frac{3\alpha_s C_F}{2\pi} , \tag{A.194}$$

$$\tilde{\gamma}_{GS}^{(\text{out})} = \frac{C_A \alpha_s}{2\pi}\log\left[\tan^2\frac{R}{2}\tan^2\frac{R_B}{2}\right] - \frac{\left(\frac{C_A}{2} - C_F\right)\alpha_s}{\pi}\log\left[\tan^2\frac{R}{2}\tan^2\frac{R_B}{2}\right] , \tag{A.195}$$

$$\gamma_{\tilde{J}_{hj}} = -4\frac{\alpha_s C_F}{2\pi(1-\alpha)}\log\left[2^{-\alpha/2}\tilde{e}_3^{(\alpha)}e^{\gamma_E}\left(\frac{\mu}{Q}\right)^\alpha N_{HJ}\right] + \frac{3\alpha_s C_F}{2\pi} , \tag{A.196}$$

$$\gamma_{\tilde{J}_{sj}} = -4\frac{\alpha_s C_A}{2\pi(1-\alpha)}\log\left[2^{-\alpha/2}\tilde{e}_3^{(\alpha)}e^{\gamma_E}z_{sj}^2\left(\frac{\mu}{z_{sj}Q}\right)^\alpha N_{SJ}\right] + \frac{\alpha_s}{2\pi}\beta_0 , \tag{A.197}$$

$$\tilde{\gamma}_{GS}^{(\text{in})} = \frac{C_F \alpha_s}{\pi(1-\alpha)}\log\left[2^{1-\alpha}\tan^{-3(1-\alpha)}\frac{R}{2}\left(\frac{2^{-\alpha}e_2^{(\alpha)}}{z_{sj}}\right)^{-3+3/\alpha}\right]$$

$$+ \frac{(C_A + 2C_F)\alpha_s}{\pi(1-\alpha)}\log\left[\frac{2^{-1+4\alpha}\mu\tilde{e}_3^{(\alpha)}e^{\gamma_E}z_{sj}}{Q}\tan^{2(1-\alpha)}\frac{R}{2}\left(\frac{2^{-\alpha}e_2^{(\alpha)}}{z_{sj}}\right)^{5-3/\alpha}\right] , \tag{A.198}$$

$$\gamma_{\tilde{S}_{n_{sj}\bar{n}_{sj}}} = \frac{C_A \alpha_s}{\pi(1-\alpha)}\log\left[\frac{\mu\tilde{e}_3^{(\alpha)}e^{\gamma_E}2^{1+2\alpha}\tan^{2(\alpha-1)}\frac{R}{2}}{Q}\left(\frac{2^{-\alpha}e_2^{(\alpha)}}{z_{sj}}\right)^{1+1/\alpha}\right] . \tag{A.199}$$

As for the soft subjet anomalous dimensions, one can check that the zero binned anomalous dimensions satisfy the consistency relation

$$
\mu\frac{d}{d\mu}\ln\tilde{H}(Q^2,\mu) + \mu\frac{d}{d\mu}\ln\tilde{H}^{sj}_{n\bar{n}}(z_{sj},n_{sj},\mu) + \mu\frac{d}{d\mu}\ln\tilde{J}_{n_{sj}}\left(\tilde{e}_3^{(\alpha)}\right) + \mu\frac{d}{d\mu}\ln\tilde{J}_{hj}\left(\tilde{e}_3^{(\alpha)}\right)
$$
$$
+ \mu\frac{d}{d\mu}\ln\tilde{J}_{oj}(R_B) + \mu\frac{d}{d\mu}\ln\tilde{S}_{n_{sj}\,\bar{n}_{sj}}\left(\tilde{e}_3^{(\alpha)};R\right) + \mu\frac{d}{d\mu}\ln\tilde{S}_{n_{sj}\,n\bar{n}}\left(\tilde{e}_3^{(\alpha)},B;R,R_B\right) = 0\,, \quad (A.200)
$$

as required for the consistency of the factorization theorem.

A.5 One Loop Calculations of Signal Factorization Theorem

In this section we give the operator definitions, and one-loop results for the functions appearing in the factorization theorem of Eq. (4.45) for the signal contribution from $Z \rightarrow q\bar{q}$. These are formulated in the SCET$_1$ effective theory of Ref. [105], in an attempt to have a consistent approach to factorization for both the signal and background distributions. In the collinear subjets region of phase space the two are identical (including identical power counting for the modes) up to the absence of global soft modes for the signal distribution. Alternatively, the factorization theorem for the signal region can be formulated by boosting the factorization theorems for appropriately chosen e^+e^- event shapes, as was considered in Ref. [85]. While this approach is less in the spirit of developing effective field theory descriptions for jet substructure that was pursued in this chapter, it has the potential advantage of being easily able to relate to higher order known results for event shapes.

Definitions of Factorized Functions

The functions appearing in the collinear subjets factorization theorem of Eq. (4.18) have the following SCET operator definitions:

- Hard Matching Coefficient:

$$H_Z\left(Q^2\right) = \left|C_Z\left(Q^2\right)\right|^2 , \qquad\qquad (A.201)$$

where $C_Z\left(Q^2\right)$ is the matrix element for the process $e^+e^- \to ZZ$, and also includes the leptonic decay of one of the Z bosons. Since we use the narrow width approximation, flat polarization distributions for the Z, and normalize our distributions to unity, it will play no role in our calculation.

- Jet Functions:

$$J_{n_{a,b}}\left(e_3^{(\alpha)}\right) = \qquad\qquad (A.202)$$
$$\frac{(2\pi)^3}{C_F}\mathrm{tr}\langle 0|\frac{\slashed{\bar{n}}_{a,b}}{2}\chi_{n_{a,b}}(0)\delta(Q - \bar{n}_{a,b}\cdot\mathcal{P})\delta^{(2)}(\vec{\mathcal{P}}_\perp)\delta\left(e_3^{(\alpha)} - \mathbf{E_3}^{(\alpha)}\right)\bar{\chi}_{n_{a,b}}(0)|0\rangle$$

- Collinear-Soft Function:

$$S_{c,\,n_a n_b}\left(e_3^{(\alpha)}\right) = \mathrm{tr}\langle 0|T\{S_{n_a}S_{n_b}\}\delta\left(e_3^{(\alpha)} - \mathbf{E_3}^{(\alpha)}\right)\bar{T}\{S_{n_a}S_{n_b}\}|0\rangle \qquad (A.203)$$

As in App. A.2 and App. A.3, the operator, $\mathbf{E_3}^{(\alpha)}$, measures the contribution to $e_3^{(\alpha)}$ from final states, and must be appropriately expanded following the power counting of the sector on which it acts. Since the power counting is identical as for the collinear subjets factorization theorem, the expansions are given in Eq. (A.24), and Eq. (A.28). In the collinear subjets region that we consider for the signal, all modes are boosted, and so there is no dependence on the jet algorithm at leading power.

Hard Matching Coefficient

The hard matching coefficient for the process $e^+e^- \to ZZ$, with one Z decaying leptonically, $H_Z(Q^2)$, does not carry an SCET anomalous dimension (hence we have dropped the μ dependence), as it is colorless. Because we work in the narrow width approximation, at a fixed

Q^2, and consider only normalized distributions, it is therefore irrelevant to our discussion.

Matrix Element for $Z \rightarrow q\bar{q}$ Decay

The anomalous dimension for the $Z \rightarrow q\bar{q}$ splitting function appearing in the factorization theorem of Eq. (4.18) is the same as that for the SCET quark bilinear operator, which was given in Eq. (A.41), but evaluated at the appropriately boosted scale.

For simplicity, in this chapter we do not account for spin correlations, and assume a flat profile in the polarization of the Z boson. The tree level $Z \rightarrow q\bar{q}$ matrix element is well known and first calculated in Ref. [486]. The full matrix element is known to two loops [487].

The anomalous dimension depends only on the color structure, and is therefore the same as the anomalous dimension for the hard matrix element for $e^+e^- \rightarrow q\bar{q}$, namely

$$\gamma_{H_Z} - 1 + \frac{\alpha_s C_F}{2\pi} \left(-8 + \frac{\pi^2}{6} - \log^2\left[\frac{\mu_H^2}{\mu^2}\right] + 3\log\left[\frac{\mu_H^2}{\mu^2}\right] \right) . \tag{A.204}$$

Here μ_H is the scale of the splitting. It is essential for the cancellation of anomalous dimensions that the scale μ_H is equal to the invariant mass of the jet. In terms of the energy correlation functions, this is given by

$$m_J^2 = \frac{Q^2 \left[z(1-z)\right]^{1-2/\alpha} \left(e_2^{(\alpha)}\right)^{2/\alpha}}{4}$$
$$= \frac{Q^2 z(1-z) n_a \cdot n_b}{2} . \tag{A.205}$$

The necessity for the appearance of the jet mass as the scale in the anomalous dimension is due to the fact that it is a Lorentz invariant quantity, and as has been discussed in Ref. [85], the factorization theorem for the case of the boosted boson can be obtained by boosting an e^+e^- event shape, where it is of course known that the scale Q^2 of the off-shell Z, or γ is the scale appearing in the hard anomalous dimension.

407

Jet Functions

The jet functions appearing in the signal factorization theorem are identical to the quark (and antiquark) jet functions calculated in App. A.2 for the collinear subjets region of phase space. This is because the power counting is identical in the two cases and the jet functions are only sensitive to the color of the jet that they describe. Therefore we do not repeat them here.

Collinear-Soft Function

The power counting for the signal is identical to the power counting for the collinear subjet region for the QCD background. However, the collinear-soft function contains only Wilson lines along the collinear subjet directions. The collinear-soft function for the QCD background was calculated in pairs of dipoles in App. A.2, and therefore the contribution from a collinear-soft exchange between the n_a and n_b Wilson lines can simply be extracted from that calculation. The result for this contribution is given by

$$
S^{(1)}_{c,\,n_a n_b}(\tilde{e}_3^{(\alpha)}) = -g^2 \Gamma(-2\epsilon) \left(\frac{\mu^2 N_{CS}^2 e^{\gamma_E} (\tilde{e}_3^{(\alpha)})^2 \left(\frac{n_a \cdot n_b}{8} \right)^{-1+2\alpha}}{4\pi 4^{1-\alpha} Q^2} \right)^{\epsilon} 16 c_\epsilon \frac{\Gamma[1/2 - \epsilon]\Gamma[1/2]}{\Gamma[1 - \epsilon]}
$$
$$
\left(\frac{1}{(2\alpha - 2)\epsilon} + \frac{\alpha \log(2)}{\alpha - 1} + \log(2) + \frac{\epsilon \left(-\pi^2 \alpha^2 + 36\alpha^2 \log^2(2) + 3\pi^2 \alpha - 24\alpha \log^2(2) - 2\pi^2 \right)}{12(\alpha - 1)} \right) ,
$$

$$
\text{(A.206)}
$$

where we recall that the normalization factor is given by

$$
N_{CS} = 2^{3\alpha/2 + 1} z_a z_b \left(n_a \cdot n_b \right)^{\alpha/2} ,
\tag{A.207}
$$

as defined in Eq. (A.74). Also note that we have factored out the color generators, so that the collinear-soft function is defined as

$$
S_c^{(1)} \left(e_3^{(\alpha)} \right) = \frac{1}{2} \sum_{i \neq j} \mathbf{T}_i \cdot \mathbf{T}_j S_{c,\,ij}^{(1)} \left(e_3^{(\alpha)} \right) ,
\tag{A.208}
$$

which is the generic form of the collinear-soft (or soft) function to one-loop.

Expanding in ϵ, and keeping only the divergent piece, as relevant for the anomalous dimensions, we find

$$
\begin{aligned}
\tilde{S}_{c,n_a n_b}^{(1)\text{div}}(\tilde{e}_3^{(\alpha)}) &= \frac{\alpha_s}{\pi}\frac{1}{(\alpha-1)\epsilon^2} + 2\frac{\alpha_s}{\pi}\frac{\left(2\alpha\log(2) + \log\left[\frac{\mu N_{CS}e^{\gamma_E}(\tilde{e}_3^{(\alpha)})\left(\frac{n_a \cdot n_b}{8}\right)^{-1/2+\alpha}}{2^{1-\alpha}Q}\right] - \log(2)\right)}{(\alpha-1)\epsilon} \\
&= \frac{\alpha_s}{\pi}\frac{1}{(\alpha-1)\epsilon^2} + 2\frac{\alpha_s}{\pi}\frac{L_\alpha^{cs}}{(\alpha-1)\epsilon},
\end{aligned} \tag{A.209}
$$

where

$$
L_\alpha^{cs} = \log\left(\frac{\mu N_{CS}e^{\gamma_E}(\tilde{e}_3^{(\alpha)})\left(n_a \cdot n_b\right)^{-1/2+\alpha}}{\sqrt{2}Q}\right). \tag{A.210}
$$

Since there is no global-soft function the cancellation of anomalous dimensions, to be discussed shortly, requires that only $1/(1-\alpha)$ contributions appear in the collinear soft function, as is observed.

Cancellation of Anomalous Dimensions

It is also interesting to explicitly check the cancellation of anomalous dimensions for the signal factorization theorem as formulated in SCET$_+$ to further confirm the cancellation mechanism which took place for the background distribution. The functions appearing in the signal factorization theorem obey identical evolution equations to those for the background distribution, which were explicitly given in App. A.2, so we do not repeat them here.

The independence of the total cross section under renormalization group evolution implies the following relation between anomalous dimensions

$$
\gamma_{H_Z} + \gamma_q^\alpha\left(\tilde{e}_3^{(\alpha)}\right) + \gamma_{\bar{q}}^\alpha\left(\tilde{e}_3^{(\alpha)}\right) + \gamma_{cs}\left(\tilde{e}_3^{(\alpha)}\right) = 0. \tag{A.211}
$$

Here γ_{H_Z} is the anomalous dimension of the $Z \to q\bar{q}$ matrix element, $\gamma_q^\alpha\left(\tilde{e}_3^{(\alpha)}\right)$ and $\gamma_{\bar{q}}^\alpha\left(\tilde{e}_3^{(\alpha)}\right)$ are the anomalous dimensions of the quark and antiquark jet functions and $\gamma_{cs}\left(\tilde{e}_3^{(\alpha)}\right)$ is the

anomalous dimension of the collinear soft function.

For the case of $Z \to q\bar{q}$, we have the color conservation relation

$$\mathbf{T}_q + \mathbf{T}_{\bar{q}} = 0\,. \tag{A.212}$$

The explicit values of the relevant Casimirs are

$$\mathbf{T}_q \cdot \mathbf{T}_q = C_F\,, \qquad \mathbf{T}_{\bar{q}} \cdot \mathbf{T}_{\bar{q}} = C_F\,, \qquad \mathbf{T}_q \cdot \mathbf{T}_{\bar{q}} = -C_F\,, \tag{A.213}$$

however, for most of the cancellation of the anomalous dimensions, it will be convenient to work in the abstract color notation.

Substituting the explicit expressions for the anomalous dimensions into the consistency relation of Eq. (A.211), we find

$$\sum_\gamma = \gamma_{H_Z} + \left[-\mathbf{T}_q \cdot \mathbf{T}_{\bar{q}} \frac{4\alpha_s L_\alpha^{cs}\left(\tilde{e}_3^{(\alpha)}\right)}{\pi(1-\alpha)} \right]$$

$$- C_F \frac{2\alpha_s L_\alpha^g\left(\tilde{e}_3^{(\alpha)}\right)}{\pi(1-\alpha)} + \gamma_q - C_F \frac{2\alpha_s L_\alpha^q\left(\tilde{e}_3^{(\alpha)}\right)}{\pi(1-\alpha)} + \gamma_q\,, \tag{A.214}$$

where $L_\alpha^g\left(\tilde{e}_3^{(\alpha)}\right)$ and $L_\alpha^q\left(\tilde{e}_3^{(\alpha)}\right)$, were defined in Eq. (A.132).

As expected, all contributions are collinear in nature, having a $1/(1-\alpha)$ dependence, and using the color conservation relation of Eq. (A.212) along with the explicit expressions for the Casimirs of Eq. (A.214), we immediately see the cancellation of the $\tilde{e}_3^{(\alpha)}$ dependence. It is also straightforward to check the cancellation of the remaining dependencies. It is a nice consistency check on the calculation that the cancellation occurs in exactly the same way as for the background cancellation, namely between the $\mathbf{T}_q \cdot \mathbf{T}_{\bar{q}}$ contribution and the jet functions. It is important to emphasize that the cancellation only occurs if the scale of the splitting is given by the invariant mass of the jet, as expected from boosting e^+e^- event shapes.

410

A.6   Soft Haze Factorization Theorem

.

For completeness, we list the operator definitions of the functions appearing in the soft haze factorization theorems. We also give the explicit forms of the measurement operators expanded in the appropriate kinematics.

The quark jet functions are given as:

$$
\begin{aligned}
J_n\!\left(e_2^{(\alpha)}\right) = \\
\frac{(2\pi)^3}{C_F}\mathrm{tr}\langle 0|\frac{\bar{\slashed{n}}_{a,b}}{2}\chi_{n_{a,b}}(0)\delta(Q-\bar{n}_{a,b}\cdot\mathcal{P})\delta^{(2)}(\vec{\mathcal{P}}_{\perp})\delta\!\left(e_2^{(\alpha)}-\mathbf{E_2}^{(\alpha)}\right)\bar{\chi}_{n_{a,b}}(0)|0\rangle\,.
\end{aligned}
\tag{A.215}
$$

The gluon jet functions are similarly defined. The soft functions appearing in the factorization theorems (4.39) and (4.40) are:

$$
\begin{aligned}
S_{n\bar{n}}\!\left(e_2^{(\beta)},e_2^{(\alpha)},e_3^{(\alpha)};R\right) = \frac{1}{C_A}\mathrm{tr}\langle 0|T\{S_nS_{\bar{n}}\}\delta\!\left(e_2^{(\beta)}-\Theta_R\mathbf{E_2}^{(\beta)}\right)\delta\!\left(e_2^{(\alpha)}-\Theta_R\mathbf{E_2}^{(\alpha)}\right) \\
\delta\!\left(e_3^{(\alpha)}-\Theta_R\mathbf{E_3}^{(\alpha)}\right)\bar{T}\{S_nS_{\bar{n}}\}|0\rangle\,,
\end{aligned}
\tag{A.216}
$$

$$
S_{n\bar{n}}\!\left(e_2^{(\alpha)},e_3^{(\alpha)};R\right) = \frac{1}{C_A}\mathrm{tr}\langle 0|T\{S_nS_{\bar{n}}\}\delta\!\left(e_2^{(\alpha)}-\Theta_R\mathbf{E_2}^{(\alpha)}\right)\delta\!\left(e_3^{(\alpha)}-\Theta_R\mathbf{E_3}^{(\alpha)}\right)\bar{T}\{S_nS_{\bar{n}}\}|0\rangle\,.
\tag{A.217}
$$

The action of the energy correlation functions on the collinear and soft haze states are given as:

$$
\mathbf{E_2}^{(\alpha)}\big|_C\big|X_n\rangle = \sum_{k,p\in X_n}\frac{\bar{n}\cdot k}{Q}\frac{\bar{n}\cdot p}{Q}\left(\frac{8p\cdot k}{\bar{n}\cdot p\,\bar{n}\cdot k}\right)^{\frac{\alpha}{2}}\big|X_n\rangle\,,
\tag{A.218}
$$

$$
\mathbf{E_2}^{(\alpha)}\big|_{SH}\big|X_s\rangle = \sum_{k\in X_s}2\frac{k^0}{Q}\left(\frac{2n\cdot k}{k^0}\right)^{\frac{\alpha}{2}}\big|X_s\rangle\,,
\tag{A.219}
$$

$$
\mathbf{E_3}^{(\alpha)}\big|_{SH}\big|X_s\rangle = \sum_{k,p\in X_s}4\frac{k^0}{Q}\frac{p^0}{Q}\left(\frac{2n\cdot k}{k^0}\frac{2n\cdot p}{p^0}\frac{2p\cdot k}{p^0k^0}\right)^{\frac{\alpha}{2}}\big|X_s\rangle\,.
\tag{A.220}
$$

411

A.7 Summary of Canonical Scales

As many of our factorization theorems involve a large number of scales, in this section we summarize for convenience the scales used in the resummation. Unless otherwise indicated, all scales are taken to be the canonical scales of the logarithms appearing in the factorization theorems.

When performing the numerical resummation, we perform the renormalization group evolution in Laplace space, and compute the cumulative distribution. We then perform the scale setting at the level of the cumulative distribution and numerically differentiate to derive the differential D_2 spectrum. While this is formally equivalent to scale setting in the differential distribution when working to all orders in perturbation theory, differences between scale setting in the differential and cumulative distribution arise when working to fixed order in perturbation theory [488]. We have not investigated the size of the effect that this has on our D_2 distributions. We utilized only two loop running of α_s, to be consistent with the Monte Carlos, and avoided the Landau pole by freezing out the running coupling at a specific $\mu_{Landau} \sim 1$ GeV.

Throughout this appendix we will use z_q and z_g to denote the energy fractions of the quark and gluon subjets, respectively. For simplicity, we restrict to the case $\alpha = \beta$. Finally, we estimate the soft out-of-jet radiation scale to be:

$$B \approx Q\left(e_2^{(\alpha)}\right)^2 \tag{A.221}$$

This is consistent with the jet algorithm constraint given by Eq. (4.26).

Collinear Subjets

We take the canonical scales for the functions appearing in the collinear subjets factorization theorem as

$$\mu_H = Q\,, \tag{A.222}$$

$$\mu_{H_2} = \frac{Q\left(e_2^{(\alpha)}\right)^{1/\alpha} z_q^{\frac{1}{2}-\frac{1}{\alpha}} z_g^{\frac{1}{2}-\frac{1}{\alpha}}}{2}, \tag{A.223}$$

$$\mu_{J_g} = \frac{e^{-\gamma_E/\alpha} Q\left(e_2^{(\alpha)}\right)^{-2/\alpha} \left(\tilde{e}_3^{(\alpha)}\right)^{1/\alpha} z_q^{\frac{1}{\alpha}} z_g}{2}, \tag{A.224}$$

$$\mu_{J_q} = \frac{e^{-\gamma_E/\alpha} Q\left(e_2^{(\alpha)}\right)^{-2/\alpha} \left(\tilde{e}_3^{(\alpha)}\right)^{1/\alpha} z_q z_g^{\frac{1}{\alpha}}}{2}, \tag{A.225}$$

$$\mu_{CS} = \frac{e^{-\gamma_E} Q\left(e_2^{(\alpha)}\right)^{-3+1/\alpha} \tilde{e}_3^{(\alpha)} z_q^{2-\frac{1}{\alpha}} z_g^{2-\frac{1}{\alpha}}}{2}, \tag{A.226}$$

$$\mu_S^{(\text{in})} = \frac{4^{-\alpha} e^{-\gamma_E} Q\left(e_2^{(\alpha)}\right)^{-1} \tilde{e}_3^{(\alpha)} \tan^2 \frac{R}{2}}{2}, \tag{A.227}$$

$$\mu_S^{(\text{out})} = B \tag{A.228}$$

where the scales are indexed by the name of the associated function in the factorization theorem.

Soft Subjets

We take the canonical scales for the functions appearing in the soft subjets factorization theorem as

$$\mu_H = Q, \tag{A.229}$$

$$\mu_{H_{sj}} = \frac{Q\left(e_2^{(\alpha)}\right)^{1/\alpha} z_{sj}^{(\alpha-1)/\alpha} \sqrt{4 - \left(e_2^{(\alpha)}\right)^{2/\alpha} z_{sj}^{-2/\alpha}}}{4}, \tag{A.230}$$

$$\mu_{S_{n_{sj}\bar{n}_{sj}}} = 2^{-\alpha} e^{-\gamma_E} Q \tan^{2(1-\alpha)} \frac{R}{2} \left(e_2^{(\alpha)}\right)^{-(1+\alpha)/\alpha} \tilde{e}_3^{(\alpha)} \left(z_{sj}\right)^{1/\alpha}$$

$$\left(1 - \frac{1}{4}\left(e_2^{(\alpha)}\right)^{2/\alpha} z_{sj}^{-2/\alpha}\right)^{(1-\alpha)/2} \left(\frac{\left(1+\tan^2\frac{R}{2}\right)\left(e_2^{(\alpha)}\right)^{2/\alpha} - 4\tan^2\frac{R}{2} z_{sj}^{2/\alpha}}{\tan^2\frac{R}{2}\left(\left(e_2^{(\alpha)}\right)^{2/\alpha} - 4z_{sj}^{2/\alpha}\right)}\right)^{1-\alpha}, \tag{A.231}$$

$$\mu_{S_{n_{sj}n\bar{n}}}^{(\text{in})} = 2^{-2+\alpha} e^{-\gamma_E} Q \tan^{2(\alpha-1)} \frac{R}{2} \left(e_2^{(\alpha)}\right)^{-5+3/\alpha} \tilde{e}_3^{(\alpha)} \left(z_{sj}\right)^{4-3/\alpha}$$

413

$$\left(1 - \frac{1}{4} \left(e_2^{(\alpha)} \right)^{2/\alpha} z_{sj}^{-2/\alpha} \right)^{(1-\alpha)/2} , \tag{A.232}$$

$$\mu_{S_{n_{sj}n\bar{n}}}^{(\text{out})} = \frac{2n \cdot n_{sj} B}{\tan \frac{\theta_{sj}}{2}} , \tag{A.233}$$

$$\mu_{J_{hj}} = \frac{Q e^{-\gamma_E/\alpha} \left(e_2^{(\alpha)} \right)^{-2/\alpha} \left(\tilde{e}_3^{(\alpha)} \right)^{1/\alpha} z_{sj}^{1/\alpha}}{2} , \tag{A.234}$$

$$\mu_{J_{\bar{n}}} = Q \tan \frac{R_B}{2} , \tag{A.235}$$

$$\mu_{J_{n_{sj}}} = \frac{Q e^{-\gamma_E/\alpha} \left(e_2^{(\alpha)} \right)^{-2/\alpha} \left(\tilde{e}_3^{(\alpha)} \right)^{1/\alpha} z_{sj}}{2} . \tag{A.236}$$

Soft Subjet Collinear Zero Bin

We take the canonical scales for the functions appearing in the collinear zero bin of the soft subjets factorization theorem as

$$\mu_{\tilde{H}} = Q , \tag{A.237}$$

$$\mu_{\tilde{H}_{sj}} = \frac{Q \left(e_2^{(\alpha)} \right)^{1/\alpha} z_{sj}^{(\alpha-1)/\alpha}}{2} , \tag{A.238}$$

$$\mu_{\tilde{S}_{n_{sj}\bar{n}_{sj}}} = \frac{2^{-\alpha} Q e^{-\gamma_E} \tan^{2(1-\alpha)} \frac{R}{2} \tilde{e}_3^{(\alpha)} z_{sj}^{1/\alpha}}{\left(e_2^{(\alpha)} \right)^{(1+\alpha)/\alpha}} , \tag{A.239}$$

$$\mu_{\tilde{S}_{n_{sj}n\bar{n}}}^{(\text{in})} = \frac{2^{-2+\alpha} Q e^{-\gamma_E} \tan^{2(\alpha-1)} \frac{R}{2} \tilde{e}_3^{(\alpha)} z_{sj}^{4-3/\alpha}}{\left(e_2^{(\alpha)} \right)^{5-3/\alpha}} , \tag{A.240}$$

$$\mu_{\tilde{S}_{n_{sj}n\bar{n}}}^{(\text{out})} = \frac{2n \cdot n_{sj} B}{\tan \frac{\theta_{sj}}{2}} , \tag{A.241}$$

$$\mu_{\tilde{J}_{hj}} = \frac{e^{-\gamma_E/\alpha} Q \left(\tilde{e}_3^{(\alpha)} \right)^{1/\alpha} z_{sj}^{1/\alpha}}{2 \left(e_2^{(\alpha)} \right)^{2/\alpha}} , \tag{A.242}$$

$$\mu_{\tilde{J}_{oj}} = Q \tan \frac{R_B}{2} , \tag{A.243}$$

$$\mu_{\tilde{J}_{n_{sj}}} = \frac{e^{-\gamma_E/\alpha} Q \left(\tilde{e}_3^{(\alpha)} \right)^{1/\alpha} z_{sj}}{2 \left(e_2^{(\alpha)} \right)^{2/\alpha}} . \tag{A.244}$$

414

Scale Variation

Here we list all the variations that went into the scale uncertainties of the QCD background calculations. Any common scale between the soft subjet factorization and its collinear bin are always varied together. Hence we will only discuss variations of the soft subjet and collinear subjets. It is important to note that $\mu_{S_{n_{sj}n\bar{n}}}^{(\text{out})}$ of the soft subjet is not exactly the same as the $\mu_S^{(\text{out})}$ of the collinear factorization. The extra angular factor improves cancellation with the soft subjet collinear zero bin in the collinear region of the phase space. In the soft subjet region, the angular factor becomes an $O(1)$ number. Given the arbitrariness of the out-of-jet scale setting, we included several different schemes.

- Splitting scales μ_{H_2} and $\mu_{H_{sj}}$ from half to twice canonical.

- μ_{Landau} where the running of the coupling is frozen from 0.5 GeV to 1.5 GeV, canonical is 1 GeV.

- All in-jet soft scales $\mu_{S_{n_{sj}n\bar{n}}}^{(\text{in})}, \mu_{\tilde{S}_{n_{sj}\bar{n}_{sj}}}, \mu_{CS}$, and μ_S from half to twice canonical. This included the scales in the collinear factorization and soft subjet factorization being varied together, and independently.

- All out-of-jet soft scales $\mu_{S_{n_{sj}n\bar{n}}}^{(\text{out})}, \mu_S^{(out)}$ from half to twice canonical. This included the scales in the collinear factorization and soft subjet factorization being varied together, and independently.

- Soft subjet out-of-jet soft scale $\mu_{S_{n_{sj}n\bar{n}}}^{(\text{out})} = Q z_{sj}^2$ from half to twice canonical. Also in this scheme the splitting scales were varied from half to twice canonical, and μ_{Landau} from 0.5 GeV to 1.5 GeV.

- Soft subjet out-of-jet soft scale $\mu_{S_{n_{sj}n\bar{n}}}^{(\text{out})} = \mu_S^{(out)}$ from half to twice canonical. Also in this scheme the splitting scales were varied from half to twice canonical, and μ_{Landau} from 0.5 GeV to 1.5 GeV.

The final uncertainty bands were taken as the envelope of these variations. Though these variations do not cover all perturbative functions that can be varied, we believe that they are representative of NLL uncertainties.

A.8 Renormalization Group Evolution of the Shape Function

In this appendix we briefly summarize some of the properties of the non-perturbative shape function used in the analysis of the D_2 observable, including hadron mass effects, so as to ensure that the level of renormalization group evolution of the parameter Ω_D is consistent with our results at both 1 TeV and 91 GeV, as discussed in Secs. 4.5.5 and 4.6, respectively. There we found that the value of Ω_D was approximately equal at the two energies, to within our uncertainties. As in the text, we assume that the dominant non-perturbative corrections arise from the global soft modes of the collinear subjets factorization theorem, so that we are working with a soft function with Wilson lines only along the n and $\bar{n}$ directions. We follow closely the formalism originally developed in Ref. [156].

In Ref. [156] it was shown that for dijet observables which can be written in terms of the rapidity y and the transverse velocity r, defined as

$$r = \frac{p_\perp}{\sqrt{p_\perp^2 + m_H^2}},$$ (A.245)

where m_H is a light hadron mass, have a leading power correction that is universal, for event shapes with the same r dependence. Furthermore, the leading power corrections can be written as an integral over an r dependent power correction,

$$\Omega_D = \int_0^1 dr \, g(r) \, \Omega_D(r),$$ (A.246)

where $g(r)$ is a function of r which depends only on the definition of the event shape (see

416

Ref. [156]), and $\Omega_D(r)$ exhibits a multiplicative renormalization group evolution in r, which is independent of y. In particular, for Ω_D, we have

$$\mu \frac{d}{d\mu}\Omega_D(r,\mu) = \gamma_{\Omega_D}(r,\mu)\Omega_D(r,\mu) = \left(-\frac{\alpha_s C_A}{\pi}\log(1-r^2)\Omega_D(r,\mu)\right),\qquad (A.247)$$

to one loop accuracy [156]. This renormalization group equation can be solved exactly for each r, however, the computation of Ω_D using Eq. (A.246) requires knowledge of the exact r dependence of $\Omega_D(r,\mu)$ at a particular scale. However, it was shown that to order α_s, only a single non-perturbative parameter is required to described the evolution, so that one can write

$$\Omega_D(\mu) = \Omega_D(\mu_0) + \frac{\alpha_s(\mu_0)C_A}{\pi}\log\left(\frac{\mu}{\mu_0}\right)\Omega_D^{\ln}(\mu_0),\qquad (A.248)$$

where apart from the non-perturbative parameter $\Omega_D(\mu_0)$ evaluated at a particular scale, we have also had to introduce the non-perturbative parameter $\Omega_D^{\ln}(\mu_0)$, which captures the logarithmic running (hence the notation).

The additional non-perturbative parameter $\Omega_D^{\ln}(\mu_0)$ is not well constrained in the literature, and therefore as a simple estimate to make sure that the values used for Ω_D at both LEP energies and at 1 TeV are consistent, we consider the estimate $\Omega_D^{\ln}(\mu_0) = \Omega_D(\mu_0)$. Making this approximation, we find the difference between the values of Ω_D as relevant for LEP and our 1 TeV analysis to differ by $\lesssim 0.1$, with the value at LEP being lower. This is small compared to our uncertainties, and compared to the scaling in the shift of the first moment with E_J and m_J. However, it is an important check that the values of Ω_D that we use are consistent with each other in our different analyses, and could be important in analyses for which jets are probed over large energy ranges.

A.9   Comparison of MC Generators for Single Emission Observables

Throughout this chapter, we have extensively compared different Monte Carlo generators both at parton and hadron level for the observable D_2, which is set by two emissions off the initiating quark. We found significant differences between different Monte Carlo generators, and as compared with our analytic calculation, particularly at parton level. After hadronization, differences remained but these were quantitative differences, not differences in the shapes of distributions. For reference, in this appendix we compare the Monte Carlo generators used in this chapter, at both parton and hadron level for an observable set by a single emission off of the initiating parton, namely the jet mass. Observables set by a single emission have been extensively studied in the literature, and are well understood. There exist automated codes for their resummation to NNLL [93, 489], and they have been extensively used to tune Monte Carlo generators. We therefore expect to see much better agreement than for the D_2 observable, demonstrating that D_2 is a more differential probe of the perturbative shower structure.[1]

In Fig. A-1 we compare the $e_2^{(2)}$ spectra both at parton and hadron level for the PYTHIA and VINCIA event generators at the Z pole. We choose to the use $e_2^{(2)}$ instead of the jet mass, as it is dimensionless. The level of agreement should be contrasted with Fig. 4-29 for the D_2 observable at the Z pole, with and without hadronization. In particular, for the $e_2^{(2)}$ observable, there is excellent agreement in the distributions at parton level, which is not true for D_2. For D_2, the disagreement is largely remedied by hadronization, while for $e_2^{(2)}$, the level of disagreement before and after hadronization is much more similar. This supports our claim that the D_2 observable provides a more differential probe of the perturbative shower in particular, and could be used to improve its description.

In Fig. A-2 we compare the $e_2^{(2)}$ spectra both at parton and hadron level for the PYTHIA,

---

[1]Differences between Monte Carlo generators for single emission observable can also be accentuated by departing from jet mass, and considering angularities, or energy correlation functions, or differences between quark and gluon jets, for which limited data from LEP can be used for tuning [37, 490].
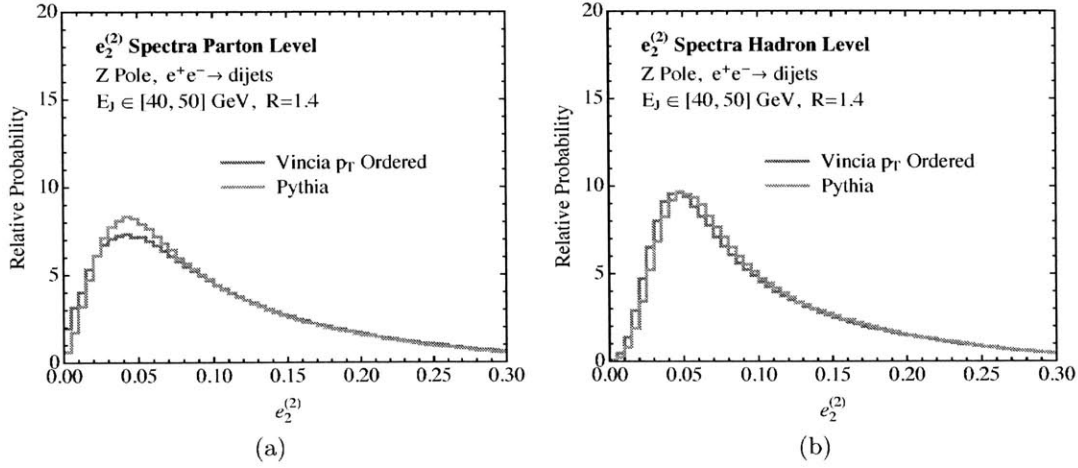
Figure A-1: A comparison of the $e_2^{(2)}$ spectrum as measured on quark initiated jets at the Z pole from the PYTHIA and p_T-ordered VINCIA Monte Carlo generators. Results are shown both for parton level Monte Carlo in a), and for hadron level Monte Carlo in b).
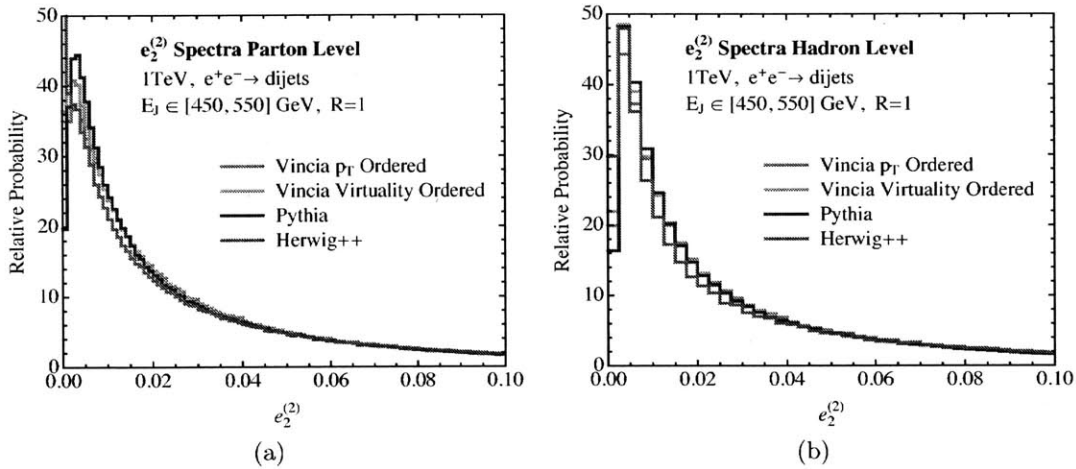
Figure A-2: A comparison of the $e_2^{(2)}$ spectrum as measured on quark initiated jets at a center of mass energy of 1 TeV from the PYTHIA, p_T-ordered VINCIA, virtuality ordered VINCIA, and HERWIG++ Monte Carlo generators. Results are shown both for parton level Monte Carlo in a), and for hadron level Monte Carlo in b).

419

p_T-ordered VINCIA, virtuality ordered VINCIA, and HERWIG++ event generators at a center of mass energy of 1 TeV and jet radius $R = 1$, as was used for the majority of numerical comparisons with analytic calculations throughout the chapter. The level of agreement in Fig. A-2 should be compared with that for the D_2 spectra throughout Sec. 4.5. In particular, it is interesting to compare the level of agreement observed for the partonic $e_2^{(2)}$ spectra as compared with the partonic D_2 spectra in Fig. 4-14. There is still some difference between the HERWIG++ spectrum at parton level and those of VINCIA and PYTHIA, however, this is to be expected, as these Monte Carlos have different hadronization models and the comparison at parton level should be taken with caution. At hadron level, all Monte Carlos also agree well for the $e_2^{(2)}$ spectra.

For completeness, in this appendix we will also include parton level plots of the $e_2^{(2)}$ distributions for the other parameter ranges that were explored in detail in the text. In Fig. A-3 we show the $e_2^{(2)}$ distributions at a center of mass energy of 500 GeV and 2 TeV, the two energies considered in the text. Only the PYTHIA and p_T-ordered VINCIA generators are considered. The level of agreement between the different generators for $e_2^{(2)}$ should be compared with the level of agreement for the D_2 spectra at these two energies, shown in Fig. 4-18. While for the D_2 observable, there was a significant discrepancy between the two generators at 2 TeV, even in the general shape of the distribution, for $e_2^{(2)}$, the distributions from the two generators agree quite well both at 500 GeV and 2 TeV. In particular, they exhibit a similar peak position and shape of the distributions.

In Fig. A-4, we consider the R dependence of the parton level $e_2^{(2)}$ distributions as measured in PYTHIA and p_T-ordered VINCIA, as was considered in Fig. 4-17 in the text for the D_2 observable. Unlike for the D_2 distributions, we see good agreement at parton level over the entire range of R. To conclude our discussion of R dependence at parton level, we also include in Fig. A-5 a comparison of the parton level D_2 spectra as measured in in PYTHIA and p_T-ordered VINCIA at 2 TeV, with $R = 0.2$ and $R = 1.0$. As was referenced in Sec. 4.5.4, while poor agreement between the two generators is seen for $R = 1$, comparably good agreement is seen at $R = 0.2$. This further supports our claim that the discrepancy

420
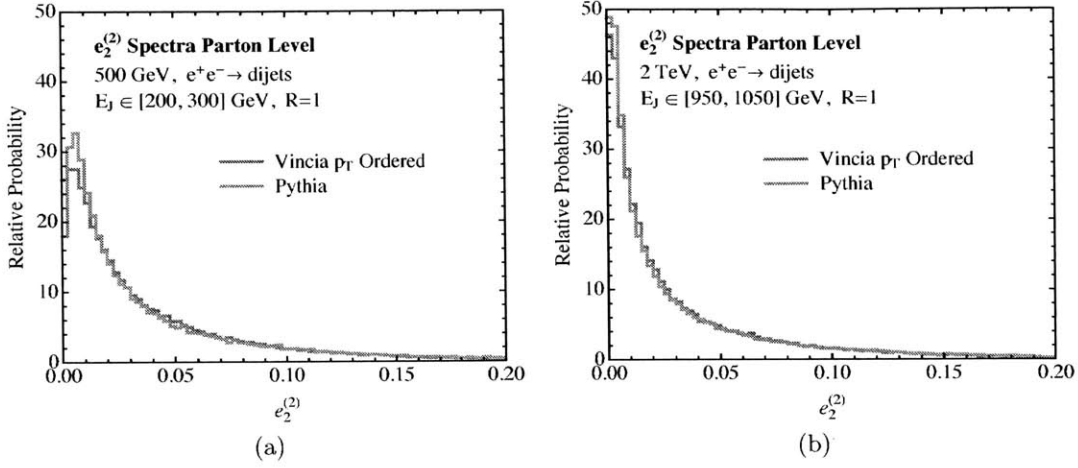
Figure A-3: A comparison of the $e_2^{(2)}$ spectrum as measured on quark initiated jets at a center of mass energy of 500 GeV in a). and 2 TeV in b). Results are shown for both the PYTHIA, and p_T-ordered VINCIA Monte Carlo generators at parton level.

between the two generators arises from the soft subjet region of phase space, which is poorly described by the PYTHIA generator, but which can be removed by considering small radius jets. We view the ability to perform analytic calculations of observables which are sensitive to the substructure of the jet in this manner as an opportunity to improve the perturbative description of the QCD shower as implemented in Monte Carlo generators.
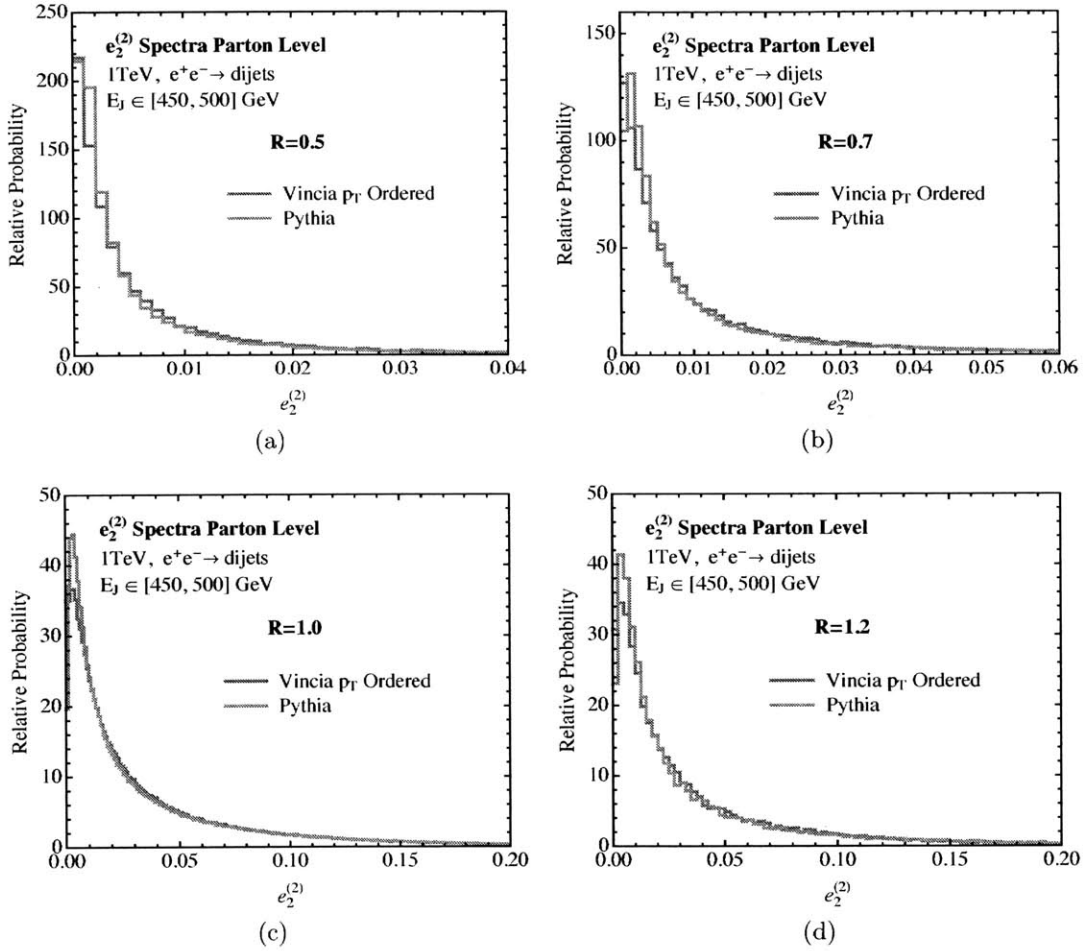
Figure A-4: A comparison of the $e_2^{(2)}$ spectrum as measured on quark initiated jets for different R values at a center of mass energy of 1 TeV from the PYTHIA, and p_T-ordered VINCIA Monte Carlo generators at parton level. Results are shown $R = 0.5, 0.7, 1.0, 1.2$ in a).-d). respectively.
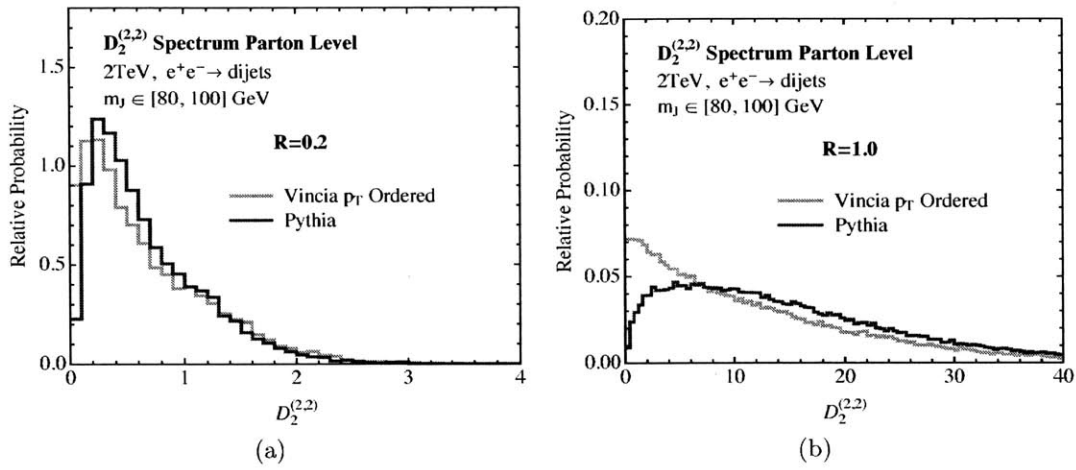
Figure A-5: A comparison of the D_2 spectrum as measured on quark initiated jets at a center of mass energy of 2 TeV from the PYTHIA, p_T-ordered VINCIA Monte Carlo generators at parton level. A jet radius of $R = 0.2$ is used in a) and $R = 1.0$ is used in b).

Appendix B

Explicit Results for Helicity Amplitudes

B.1 Spinor and Color Identities

B.1.1 Spinor Algebra

The overall phase of the spinors $|p\pm\rangle$ is not determined by the Dirac equation, $\not{p} \, |p\pm\rangle = 0$, and so can be chosen freely. In the Dirac representation,

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad \gamma_5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{B.1}$$

and taking $n_i^\mu = (1, 0, 0, 1)$, we have the standard solutions [331]

$$|p+\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{p^-} \\ \sqrt{p^+}e^{i\phi_p} \\ \sqrt{p^-} \\ \sqrt{p^+}e^{i\phi_p} \end{pmatrix}, \quad |p-\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{p^+}e^{-i\phi_p} \\ -\sqrt{p^-} \\ -\sqrt{p^+}e^{-i\phi_p} \\ \sqrt{p^-} \end{pmatrix}, \tag{B.2}$$

where

$$p^\pm = p^0 \mp p^3, \qquad \exp(\pm i\phi_p) = \frac{p^1 \pm ip^2}{\sqrt{p^+ p^-}}. \tag{B.3}$$

425

For negative p^0 and $p^\pm$ we use the usual branch of the square root, such that for $p^0 > 0$

$$|(-p)\pm\rangle = \mathrm{i}\,|p\pm\rangle \ . \tag{B.4}$$

The conjugate spinors, $\langle p\pm|$, are defined as

$$\langle p\pm| = \mathrm{sgn}(p^0)\,\overline{|p\pm\rangle} \ . \tag{B.5}$$

The additional minus sign for negative p^0 is included to use the same branch of the square root for both types of spinors, i.e., for $p^0 > 0$

$$\langle(-p)\pm| = -\overline{|(-p)\pm\rangle} = -(-\mathrm{i})\,\langle p\pm| = \mathrm{i}\,\langle p\pm| \ . \tag{B.6}$$

In this way all spinor identities are automatically valid for both positive and negative momenta, which makes it easy to use crossing symmetry. The additional signs only appear in relations which involve explicit complex conjugation. The most relevant is

$$\langle p-\rangle\, q+^* = \mathrm{sgn}(p^0 q^0)\,\langle q+\rangle\, p- \ . \tag{B.7}$$

The spinor products are denoted by

$$\langle pq\rangle = \langle p-\rangle\, q+ \ , \qquad [pq] = \langle p+\rangle\, q- \ . \tag{B.8}$$

Similarly, for products involving additional gamma matrices, we write

$$\langle p|\,\gamma^\mu|q] = \langle p-|\,\gamma^\mu\,|q-\rangle \ , \quad [p|\gamma^\mu\,|q\rangle = \langle p+|\,\gamma^\mu\,|q+\rangle \ , \tag{B.9}$$

$$\langle p|\,k|q] = \langle p-|\,\not{k}\,|q-\rangle \ , \quad [p|k\,|q\rangle = \langle p+|\,\not{k}\,|q+\rangle \ , \tag{B.10}$$

$$\langle p|\,qk\,|l\rangle = \langle p-|\,\not{q}\not{k}\,|l+\rangle \ , \quad [p|qk|l] = \langle p+|\,\not{q}\not{k}\,|l-\rangle \ , \tag{B.11}$$

etc.

426

Some useful identities, that follow directly from the definition of the spinors, are

$$\langle pq \rangle = -\langle qp \rangle, \qquad [pq] = -[qp], \tag{B.12}$$

$$[p|\gamma^\mu|p\rangle = \langle p|\gamma^\mu|p] = 2p^\mu. \tag{B.13}$$

From the completeness relations

$$|p\pm\rangle \langle p\pm| = \frac{1 \pm \gamma_5}{2} \not{p}, \tag{B.14}$$

$$\not{p} = |p] \langle p| + |p\rangle [p|, \tag{B.15}$$

one finds

$$\langle pq \rangle [qp] = \frac{1}{2} \operatorname{tr}\{(1 - \gamma_5)\not{p}\not{q}\} = 2p \cdot q. \tag{B.16}$$

Combining this with Eq. (B.7), it follows that

$$|\langle pq \rangle| = |[pq]| = \sqrt{|2p \cdot q|}. \tag{B.17}$$

The completeness relation is also useful to reduce typical expressions like

$$[p|q\,|k\rangle = [pq]\langle qk \rangle, \tag{B.18}$$

to spinor products.

Charge conjugation invariance of the current, the Fierz identity and the Schouten identity are

$$\langle p|\gamma^\mu|q] = [q|\gamma^\mu|p\rangle, \tag{B.19}$$

$$[p|\gamma_\mu|q\rangle [k|\gamma^\mu|l\rangle = 2[pk]\langle lq \rangle, \tag{B.20}$$

$$\langle pq \rangle \langle kl \rangle = \langle pk \rangle \langle ql \rangle + \langle pl \rangle \langle kq \rangle. \tag{B.21}$$

Finally, momentum conservation $\sum_{i=1}^{n} p_i = 0$ implies

$$\sum_{i=1}^{n} [ji]\langle ik \rangle = 0 . \qquad (B.22)$$

From Eq. (B.2), we see that under parity the spinors transform as

$$|p^{\mathrm{P}}\pm\rangle = \pm e^{\pm i\phi_p}\gamma^0 |p\mp\rangle , \qquad (B.23)$$

and therefore

$$\langle p^{\mathrm{P}} q^{\mathrm{P}} \rangle = -e^{i(\phi_p+\phi_q)}[pq] , \qquad (B.24)$$

$$[p^{\mathrm{P}} q^{\mathrm{P}}] = -e^{-i(\phi_p+\phi_q)}\langle pq \rangle . \qquad (B.25)$$

When applying the above result to a helicity amplitude, the phases which appear are determined by the little group scaling (see e.g. Refs. [331, 332, 491] for a review). The little group is the subgroup of the Lorentz transformations that fixes a particular momentum. In terms of the spinor helicity variables, the action of the little group, which preserves the momentum vector p, is given by

$$|p\rangle \rightarrow z|p\rangle, \qquad [p| \rightarrow \frac{1}{z}[p| . \qquad (B.26)$$

In the case that the particle with momentum p has helicity h, the corresponding helicity amplitude scales as z^{-2h} under the little group scaling. This property of the helicity amplitudes then predicts the phases that appear in the amplitude under a parity transformation.

The following completeness relation for the polarization vectors is also useful

$$\sum_{\lambda=\pm} \epsilon_\mu^\lambda(p,q) \left(\epsilon_\nu^\lambda(p,q)\right)^* = -g_{\mu\nu} + \frac{p_\mu q_\nu + p_\nu q_\mu}{p \cdot q} . \qquad (B.27)$$

428

In SCET the collinear quark fields produce projected spinors

$$|p\pm\rangle_n = \frac{\not{n}\not{\bar{n}}}{4}\,|p\pm\rangle\,.\tag{B.28}$$

The projected spinor trivially satisfies the relation

$$\not{n}\left(\frac{\not{n}\not{\bar{n}}}{4}\,|p\pm\rangle\right) = 0\,,\tag{B.29}$$

so it is proportional to $|n\pm\rangle$. Working in the basis in Eq. (B.2), we have

$$
\begin{aligned}
\frac{\not{n}\not{\bar{n}}}{4}\,|p\rangle =& \sqrt{p^0}\left[\cos\left(\frac{\theta_n}{2}\right)\cos\left(\frac{\theta_p}{2}\right)\right.\\
&\left.+e^{i(\phi_p-\phi_n)}\sin\left(\frac{\theta_n}{2}\right)\sin\left(\frac{\theta_p}{2}\right)\right]|n\rangle\,,\\
\frac{\not{n}\not{\bar{n}}}{4}\,|p] =& \sqrt{p^0}\left[e^{i(\phi_p-\phi_n)}\cos\left(\frac{\theta_n}{2}\right)\cos\left(\frac{\theta_p}{2}\right)\right.\\
&\left.+\sin\left(\frac{\theta_n}{2}\right)\sin\left(\frac{\theta_p}{2}\right)\right]|n]\,.
\end{aligned}\tag{B.30}
$$

Here θ_n, ϕ_n, and θ_p, ϕ_p, are the polar and azimuthal angle of the n and p vectors, respectively. In particular, we see that choosing $n^\mu = p^\mu/p^0$, which can always be done at leading power since there is a single particle per collinear sector, we have $\phi_p = \phi_n$, $\theta_p = \theta_n$, and the simple relation

$$\frac{\not{n}\not{\bar{n}}}{4}\,|p\pm\rangle = \sqrt{\frac{\mathbf{n}\cdot p}{2}}\,|n\pm\rangle\,.\tag{B.31}$$

B.1.2   Color Algebra

The generators t_r^a of a general irreducible representation r of SU(N) satisfy

$$[t_r^a, t_r^b] = i f^{abc}\,t_r^c\,,\quad t_r^a\,t_r^a = C_r\mathbf{1}\,,\quad \mathrm{tr}[t_r^a\,t_r^b] = T_r\,\delta^{ab}\,,\tag{B.32}$$

429

where f^{abc} are completely antisymmetric, and C_r is the quadratic Casimir of the representation r. The normalization T_r is given by $T_r = C_r d_r/d$, where d_r is the dimension of the representation and d the dimension of the Lie algebra.

We denote the generators in the fundamental representation by $t_F^a = T^a$, and the overall normalization is fixed by choosing a specific value for T_F. The adjoint representation is given by $(t_A^a)_{bc} = -if^{abc}$, which implies

$$f^{acd}f^{bcd} = C_A\,\delta^{ab}\,.\tag{B.33}$$

We also define the symmetric structure constants as

$$d^{abc} = \frac{1}{T_F}\mathrm{tr}[T^a\{T^b, T^c\}]\,.\tag{B.34}$$

For the fundamental and adjoint representations we have $d_F = N$, $d_A = d = N^2 - 1$, and so

$$C_F = \frac{N^2 - 1}{2N}\,,\qquad C_A = N\,,\tag{B.35}$$

where we have chosen the standard normalization

$$T_F = \frac{1}{2}\,.\tag{B.36}$$

Throughout the text, and for the amplitudes in the appendices, we have kept T_F arbitrary. This can be done using $C_F = T_F(N^2 - 1)/N$, $C_A = 2T_F N$. The strong coupling constant, g_s, can be kept convention independent, by using $g_s \to g_s/\sqrt{2T_F}$.

Some additional useful color identities are

$$t_r^a t_r^b t_r^a = \left(C_r - \frac{C_A}{2}\right) t_r^b\,,\tag{B.37}$$

$$T^a T^b T^c T^a = T_F^2\,\delta^{bc}\mathbf{1} + \left(C_F - \frac{C_A}{2}\right) T^b T^c\,,\tag{B.38}$$

430

where the second relation is equivalent to the completeness relation

$$
T_{\alpha\bar\beta}^a \, T_{\gamma\bar\delta}^a = T_F\left(\delta_{\alpha\bar\delta}\,\delta_{\gamma\bar\beta} - \frac{1}{N}\delta_{\alpha\bar\beta}\,\delta_{\gamma\bar\delta}\right).
\tag{B.39}
$$

We also have

$$
T^b\,\mathrm{i}f^{bac}\,T^c = \frac{C_A}{2}\,T^a\,,
\tag{B.40}
$$

$$
T^c\,\mathrm{i}f^{cad}\mathrm{i}f^{dbe}T^e = T_F^2\,\delta^{ab}\,\mathbf{1} + \frac{C_A}{2}\,T^aT^b\,.
\tag{B.41}
$$

B.1.3 QCD Color Decompositions

Here we briefly review a common color decomposition for QCD NLO amplitudes [226, 227, 228, 229]. The color bases used for the processes discussed in the text are specific examples of the decompositions given below, and were chosen to facilitate the extraction of the matching coefficients from the amplitudes literature. For a pedagogical introduction to color decompositions in QCD amplitudes see for example Refs. [331, 332].

For an n gluon process, a one-loop color decomposition in terms of fundamental generators T^a is given by

$$
\begin{aligned}
\mathcal{A}_n(g_1\cdots g_n) &= g_s^{n-2}\sum_{\sigma\in S_n/Z_n}\mathrm{tr}[T^{a_{\sigma(1)}}\cdots T^{a_{\sigma(n)}}]\Big[A_n^{\mathrm{tree}}\big(\sigma(1),\cdots,\sigma(n)\big) + g_s^2\,C_A A_{n;1}\big(\sigma(1),\cdots,\sigma(n)\big)\Big]\\
&\quad + g_s^n\sum_{c=3}^{\lfloor n/2\rfloor+1}\sum_{\sigma\in S_n/S_{c-1,n-c+1}}\mathrm{tr}[T^{a_{\sigma(1)}}\cdots T^{a_{\sigma(c-1)}}]\mathrm{tr}[T^{a_{\sigma(c)}}\cdots T^{a_{\sigma(n)}}]A_{n;c}\big(\sigma(1),\cdots,\sigma(n)\big)\,,
\end{aligned}
\tag{B.42}
$$

where $A_{n;1}$, A_n^{tree} are primitive amplitudes, which can be efficiently calculated using unitarity methods, and the $A_{n;c}$ are partial amplitudes which can be written as sums of permutations of the primitive amplitudes. The amplitudes appearing in this decomposition are separately gauge invariant. In this formula, S_n is the permutation group on n elements, and $S_{i,j}$ is the subgroup of S_{i+j} which leaves the given trace structure invariant. At tree level, only the

431

single trace color structure appears.

In the case that additional noncolored particles are also present, an identical decomposition exists, since the color structure is unaffected. For example, for a process involving n gluons and a Higgs particle, the amplitude satisfies the same decomposition as in Eq. (B.42), but with the partial and primitive amplitudes in Eq. (B.42) simply replaced by $A\big(\phi, \sigma(1), \cdots, \sigma(n)\big)$, where ϕ denotes the Higgs particle [358].

A similar decomposition exists for processes involving $q\bar{q}$ pairs. For example, the one-loop decomposition for a process with a $q\bar{q}$ pair and $n - 2$ gluons is given by [246]

$$
\begin{aligned}
&\mathcal{A}_n\big(\bar{q}_1 q_2 g_3 \ldots g_n\big) \\
&= g_s^{n-2} \sum_{\sigma \in S_{n-2}} \big(T^{a_{\sigma(3)}} \cdots T^{a_{\sigma(n)}}\big)_{\alpha\bar{\beta}} \Big[A_n^{\text{tree}}\big(1_{\bar{q}}, 2_q; \sigma(3), \cdots, \sigma(n)\big) + g_s^2\, C_A A_{n;1}\big(1_{\bar{q}}, 2_q; \sigma(3), \cdots, \sigma(n)\big) \Big] \\
&+ g_s^n \sum_{c=3}^{n-3} \sum_{\sigma \in S_{n-2}/Z_{c-1}} \text{tr}[T^{a_{\sigma(3)}} \cdots T^{a_{\sigma(c+1)}}]\big(T^{a_{\sigma(c+2)}} \cdots T^{a_{\sigma(n)}}\big)_{\alpha\bar{\beta}} A_{n;c}\big(1_{\bar{q}}, 2_q; \sigma(3), \cdots, \sigma(n)\big) \\
&+ g_s^n \sum_{\sigma \in S_{n-2}/Z_{n-3}} \text{tr}[T^{a_{\sigma(3)}} \cdots T^{a_{\sigma(n-1)}}]\big(T^{a_{\sigma(n)}}\big)_{\alpha\bar{\beta}} A_{n;n-2}\big(1_{\bar{q}}, 2_q; \sigma(3), \cdots, \sigma(n)\big) \\
&+ g_s^n \sum_{\sigma \in S_{n-2}/Z_{n-2}} \text{tr}[T^{a_{\sigma(3)}} \cdots T^{a_{\sigma(n)}}]\delta_{\alpha\bar{\beta}} A_{n;n-1}\big(1_{\bar{q}}, 2_q; \sigma(3), \cdots, \sigma(n)\big) \, .
\end{aligned}
\tag{B.43}
$$

This decomposition is easily extended to the case of additional $q\bar{q}$ pairs. As with the gluon case, the same color decomposition also applies if additional uncolored particles are included in the amplitude.

For more than five particles, the one-loop color decompositions given above do not give a complete basis of color structures beyond one loop, since color structures with more than two traces can appear. A complete basis of color structures is required for the SCET basis to guarantee a consistent RGE. A convenient basis of color structures for one-loop matching is then given by extending the one-loop decomposition to involve all higher trace structures.

432

B.2 Helicity Amplitudes for Higgs + Jets

In this appendix we give explicit results for the hard matching coefficients for $H + 0, 1, 2$ jets. We only explicitly consider gluon-fusion processes, where the Higgs couples to two gluons through a top-quark loop, and additional jets correspond to additional gluons, or quark antiquark pairs. When matching onto SCET we perform a one-step matching and directly match full QCD onto SCET, as was done for $H + 0$ jets in Ref. [300]. Most QCD results are obtained in the limit of infinite top quark mass, by first integrating out the top quark and matching onto an effective ggH interaction,

$$\mathcal{L}_{\text{hard}} = \frac{C_1}{12\pi v} H G^a_{\mu\nu} G^{\mu\nu a} \,, \tag{B.44}$$

which is then used to compute the QCD amplitudes. Here $v = (\sqrt{2}G_F)^{-1/2} = 246\,\text{GeV}$. From the point of view of the one-step matching from QCD onto SCET, using Eq. (B.44) is just a convenient way to compute the full QCD amplitude in the $m_t \to \infty$ limit. In particular, the α_s corrections to C_1 in Eq. (B.44) are included in the amplitudes below, and therefore also in the SCET Wilson coefficients. In this way, if higher-order corrections in $1/m_t$ or the exact m_t dependence for a specific amplitude are known, they can easily be included in the QCD amplitudes and the corresponding SCET Wilson coefficients. We illustrate this for the case of $H + 0$ jets below.

We separate the QCD amplitudes into their IR-divergent and IR-finite parts

$$A = A_{\text{div}} + A_{\text{fin}} \,,$$
$$B = B_{\text{div}} + B_{\text{fin}} \,, \tag{B.45}$$

where A_{fin}, B_{fin} enter the matching coefficients in Sec. 5.5. For simplicity, we drop the subscript "fin" for those amplitudes that have no divergent parts, i.e. for $A_{\text{div}} = 0$ we have

433

$A_{\text{fin}} \equiv A$. For the logarithms we use the notation

$$L_{ij} = \ln\left(-\frac{s_{ij}}{\mu^2} - i0\right), \qquad L_{ij/H} = \ln\left(-\frac{s_{ij}}{\mu^2} - i0\right) - \ln\left(-\frac{m_H^2}{\mu^2} - i0\right).$$

B.2.1 $H + 0$ Jets

We expand the amplitudes in powers of $\alpha_s(\mu)$ as

$$A = \frac{2T_F \alpha_s(\mu)}{3\pi v} \sum_{n=0}^{\infty} A^{(n)} \left(\frac{\alpha_s(\mu)}{4\pi}\right)^n. \tag{B.46}$$

The amplitudes with opposite helicity gluons vanish to all orders because of angular momentum conservation,

$$A(1^{\pm}, 2^{\mp}; 3_H) = 0, \tag{B.47}$$

corresponding to the fact that the helicity operators for these helicity configurations were not included in the basis of Eq. (5.104). The lowest order helicity amplitudes including the full m_t dependence are given by

$$A^{(0)}(1^+, 2^+; 3_H) = \frac{s_{12}}{2} \frac{[12]}{\langle 12 \rangle} F^{(0)}\left(\frac{s_{12}}{4m_t^2}\right) = \frac{s_{12}}{2} F^{(0)}\left(\frac{s_{12}}{4m_t^2}\right) e^{i\Phi_{++H}},$$

$$A^{(0)}(1^-, 2^-; 3_H) = \frac{s_{12}}{2} \frac{\langle 12 \rangle}{[12]} F^{(0)}\left(\frac{s_{12}}{4m_t^2}\right) = \frac{s_{12}}{2} F^{(0)}\left(\frac{s_{12}}{4m_t^2}\right) e^{i\Phi_{--H}}, \tag{B.48}$$

where the function $F^{(0)}(z)$ is defined as

$$F^{(0)}(z) = \frac{3}{2z} - \frac{3}{2z}\left|1 - \frac{1}{z}\right| \begin{cases} \arcsin^2(\sqrt{z}), & 0 < z \leq 1, \\ \ln^2[-i(\sqrt{z} + \sqrt{z-1})], & z > 1. \end{cases} \tag{B.49}$$

For simplicity, we have extracted the (irrelevant) overall phases

$$e^{i\Phi_{++H}} = \frac{[12]}{\langle 12 \rangle}, \qquad e^{i\Phi_{--H}} = \frac{\langle 12 \rangle}{[12]}. \tag{B.50}$$

Since the two helicity amplitudes for ggH cannot interfere and are equal to each other by parity up to an overall phase, their higher-order corrections are the same as for the spin-summed ggH form factor. The divergent part of the NLO amplitudes is given by

$$A_{\text{div}}^{(1)}(1^{\pm}, 2^{\pm}; 3_H) = A^{(0)}(1^{\pm}, 2^{\pm}; 3_H)\left[-\frac{2}{\epsilon^2}C_A + \frac{1}{\epsilon}(2C_A L_{12} - \beta_0)\right]. \tag{B.51}$$

The IR-finite parts entering the matching coefficients in Eq. (5.107) at NLO are [300]

$$
\begin{aligned}
A_{\text{fin}}^{(1)}(1^{\pm}, 2^{\pm}; 3_H) &= A^{(0)}(1^{\pm}, 2^{\pm}; 3_H)\left[C_A\left(-L_{12}^2 + \frac{\pi^2}{6}\right) + F^{(1)}\left(\frac{s_{12}}{4m_t^2}\right)\right], \\
F^{(1)}(z) &= C_A\left(5 - \frac{38}{45}z - \frac{1289}{4725}z^2 - \frac{155}{1134}z^3 - \frac{5385047}{65488500}z^4\right) \\
&\quad + C_F\left(-3 + \frac{307}{90}z + \frac{25813}{18900}z^2 + \frac{3055907}{3969000}z^3 + \frac{659504801}{1309770000}z^4\right) + \mathcal{O}(z^5).
\end{aligned}
\tag{B.52}
$$

The full analytic expression for $F^{(1)}(z)$ is very long, so we only give the result expanded in z. Since the additional m_t dependence coming from $F^{(1)}(z)$ is small and the expansion converges quickly, the expanded result is fully sufficient for on-shell studies of Higgs production. The IR-finite parts at NNLO are [300]

$$
\begin{aligned}
&A_{\text{fin}}^{(2)}(1^{\pm}, 2^{\pm}; 3_H) \\
&= A^{(0)}(1^{\pm}, 2^{\pm}; 3_H)\left\{\frac{1}{2}C_A^2 L_{12}^4 + \frac{1}{3}C_A\beta_0 L_{12}^3 + C_A\left[C_A\left(-\frac{4}{3} + \frac{\pi^2}{6}\right) - \frac{5}{3}\beta_0 - F^{(1)}\left(\frac{s_{12}}{4m_t^2}\right)\right]L_{12}^2 \right. \\
&\quad \left. + \left[C_A^2\left(\frac{59}{9} - 2\zeta_3\right) + C_A\beta_0\left(\frac{19}{9} - \frac{\pi^2}{3}\right) - \beta_0 F^{(1)}\left(\frac{s_{12}}{4m_t^2}\right)\right]L_{12} + F^{(2)}\left(\frac{s_{12}}{4m_t^2}\right)\right\}, \\
F^{(2)}(z) &= \left(7C_A^2 + 11C_A C_F - 6C_F\beta_0\right)\ln(-4z - i0) + C_A^2\left(-\frac{419}{27} + \frac{7\pi^2}{6} + \frac{\pi^4}{72} - 44\zeta_3\right) \\
&\quad + C_A C_F\left(-\frac{217}{2} - \frac{\pi^2}{2} + 44\zeta_3\right) + C_A\beta_0\left(\frac{2255}{108} + \frac{5\pi^2}{12} + \frac{23\zeta_3}{3}\right) - \frac{5}{6}C_A T_F \\
&\quad + \frac{27}{2}C_F^2 + C_F\beta_0\left(\frac{41}{2} - 12\zeta_3\right) - \frac{4}{3}C_F T_F + \mathcal{O}(z).
\end{aligned}
\tag{B.53}
$$

Here we only give the leading terms in the $m_t \to \infty$ limit. The first few higher-order terms in z in $F^{(2)}(z)$ can be obtained from the results of Refs. [352, 353].

B.2.2 $H + 1$ Jet

The amplitudes for $H + 1$ jet were calculated in Ref. [252] in the $m_t \to \infty$ limit. Reference [252] uses $T_F = 1$ and $g_s T^a / \sqrt{2}$ for the $q\bar{q}g$ coupling. Thus, we can convert to our conventions by replacing $T^a \to \sqrt{2} T^a$, and identifying $1/N = C_A - 2C_F$ and $N = C_A$ in the results of Ref. [252]. We expand the amplitudes in powers of $\alpha_s(\mu)$ as

$$A = \frac{2 T_F \alpha_s(\mu)}{3\pi v} g_s(\mu) \sum_{n=0}^{\infty} A^{(n)} \left(\frac{\alpha_s(\mu)}{4\pi} \right)^n . \tag{B.54}$$

$gggH$

The tree-level amplitudes entering the matching coefficient $\vec{C}_{++\pm}$ in Eq. (5.115) are

$$A^{(0)}(1^+, 2^+, 3^+; 4_H) = \frac{1}{\sqrt{2}} \frac{m_H^4}{\langle 12 \rangle \langle 23 \rangle \langle 31 \rangle} = \frac{m_H^4}{\sqrt{2|s_{12}s_{13}s_{23}|}} e^{i\Phi_{+++H}} ,$$

$$A^{(0)}(1^+, 2^+, 3^-; 4_H) = \frac{1}{\sqrt{2}} \frac{[12]^3}{[13][23]} = \frac{s_{12}^2}{\sqrt{2|s_{12}s_{13}s_{23}|}} e^{i\Phi_{++-H}} , \tag{B.55}$$

where we have extracted the (irrelevant) overall phases

$$e^{i\Phi_{+++H}} = \frac{\sqrt{|s_{12}|}}{\langle 12 \rangle} \frac{\sqrt{|s_{13}|}}{\langle 31 \rangle} \frac{\sqrt{|s_{23}|}}{\langle 23 \rangle} , \qquad e^{i\Phi_{++-H}} = \frac{[12]}{\langle 12 \rangle} \frac{\sqrt{|s_{12}|}}{\langle 12 \rangle} \frac{\sqrt{|s_{13}|}}{[13]} \frac{\sqrt{|s_{23}|}}{[23]} . \tag{B.56}$$

The divergent parts of the one-loop amplitudes are

$$A^{(1)}_{\text{div}}(1^+, 2^+, 3^\pm; 4_H) = A^{(0)}(1^+, 2^+, 3^\pm, 4_H) \left\{ -\frac{3}{\epsilon^2} C_A + \frac{1}{\epsilon} \left[C_A (L_{12} + L_{13} + L_{23}) - \frac{3}{2}\beta_0 \right] \right\} . \tag{B.57}$$

The finite parts of the $gggH$ amplitudes, which enter the matching coefficient $\vec{C}_{++\pm}$ at one loop are

$$A^{(1)}_{\text{fin}}(1^+, 2^+, 3^+; 4_H)$$

$$= A^{(0)}(1^+, 2^+, 3^+; 4_H) \left\{ f(s_{12}, s_{13}, s_{23}, m_H^2, \mu) + \frac{1}{3}(C_A - 2T_F n_f) \frac{s_{12}s_{13} + s_{12}s_{23} + s_{13}s_{23}}{m_H^4} \right\} ,$$

$$A_{\text{fin}}^{(1)}(1^+,2^+,3^-;4_H) = A^{(0)}(1^+,2^+,3^-;4_H)\left\{ f(s_{12},s_{13},s_{23},m_H^2,\mu) + \frac{1}{3}(C_A - 2T_F n_f)\frac{s_{13}s_{23}}{s_{12}^2} \right\},$$

$$\text{(B.58)}$$

where we have extracted the common function

$$f(s_{12},s_{13},s_{23},m_H^2,\mu) = -C_A\left[\frac{1}{2}(L_{12}^2 + L_{13}^2 + L_{23}^2) + L_{12/H}L_{13/H} + L_{12/H}L_{23/H} + L_{13/H}L_{23/H}\right.$$

$$\left. + 2\text{Li}_2\left(1 - \frac{s_{12}}{m_H^2}\right) + 2\text{Li}_2\left(1 - \frac{s_{13}}{m_H^2}\right) + 2\text{Li}_2\left(1 - \frac{s_{23}}{m_H^2}\right) - 5 - \frac{3\pi^2}{4}\right] - 3C_F.$$

$$\text{(B.59)}$$

$g q \bar{q} H$

The tree-level amplitudes entering the matching coefficient $\vec{C}_{\pm(+)}$ in Eq. (5.114) are

$$A^{(0)}(1^+;2_q^+,3_{\bar{q}}^-;4_H) = -\frac{1}{\sqrt{2}}\frac{[12]^2}{[23]} = \frac{s_{12}}{\sqrt{2|s_{23}|}}e^{i\Phi_{+(+)H}},$$

$$A^{(0)}(1^-;2_q^+,3_{\bar{q}}^-;4_H) = -\frac{1}{\sqrt{2}}\frac{\langle 13\rangle^2}{\langle 23\rangle} = \frac{s_{13}}{\sqrt{2|s_{23}|}}e^{i\Phi_{-(+)H}}, \qquad \text{(B.60)}$$

where the (irrelevant) overall phases are given by

$$e^{i\Phi_{+(+)H}} = \frac{[12]}{\langle 12\rangle}\frac{\sqrt{|s_{23}|}}{[23]}, \qquad e^{i\Phi_{-(+)H}} = \frac{\langle 13\rangle}{[13]}\frac{\sqrt{|s_{23}|}}{\langle 23\rangle}. \qquad \text{(B.61)}$$

The divergent parts of the one-loop amplitudes are

$$A_{\text{div}}^{(1)}(1^\pm;2_q^+,3_{\bar{q}}^-;4_H) = A^{(0)}(1^\pm;2_q^+,3_{\bar{q}}^-;4_H)\left\{-\frac{1}{\epsilon^2}(C_A+2C_F)+\frac{1}{\epsilon}\left[C_A(L_{12}+L_{13}-L_{23})+C_F(2L_{23}-3)-\frac{\beta_0}{2}\right]\right\}.$$

$$\text{(B.62)}$$

The finite parts of the $g q \bar{q} H$ amplitudes, which enter the matching coefficient $\vec{C}_{\pm(+)}$ at one loop are

$$A_{\text{fin}}^{(1)}(1^+;2_q^+,3_{\bar{q}}^-;4_H) = A^{(0)}(1^+;2_q^+,3_{\bar{q}}^-;4_H)\left\{g(s_{12},s_{13},s_{23},m_H^2,\mu) + (C_F - C_A)\frac{s_{23}}{s_{12}}\right\},$$

$$A_{\text{fin}}^{(1)}(1^-; 2_q^+, 3_{\bar{q}}^-; 4_H) = A^{(0)}(1^-; 2_q^+, 3_{\bar{q}}^-; 4_H)\left\{ g(s_{12}, s_{13}, s_{23}, m_H^2, \mu) + (C_F - C_A)\frac{s_{23}}{s_{13}} \right\},$$

$$(B.63)$$

where we have extracted the common function

$$
\begin{aligned}
&g(s_{12}, s_{13}, s_{23}, m_H^2, \mu) \\
&= C_A\left[-\frac{1}{2}(L_{12}^2 + L_{13}^2 - L_{23}^2) + L_{12/H}L_{13/H} - (L_{12/H} + L_{13/H})L_{23/H} - 2\text{Li}_2\left(1 - \frac{s_{23}}{m_H^2}\right) \right. \\
&\quad \left. + \frac{22}{3} + \frac{\pi^2}{4} \right] + C_F\left[-L_{23}^2 + 3L_{23} - 2L_{12/H}L_{13/H} - 2\text{Li}_2\left(1 - \frac{s_{12}}{m_H^2}\right) - 2\text{Li}_2\left(1 - \frac{s_{13}}{m_H^2}\right) \right. \\
&\quad \left. - 11 + \frac{\pi^2}{2} \right] + \beta_0\left(-L_{23} + \frac{5}{3} \right).
\end{aligned}
$$

$$(B.64)$$

B.2.3 $H + 2$ Jets

The full set of tree-level helicity amplitudes for $H+2$ jets in the $m_t \to \infty$ limit were calculated in Ref. [251], and all amplitudes below are taken from there. We expand the amplitudes A, B, in the decomposition of Eq. (5.119), Eq. (5.124), and Eq. (5.130), as

$$
\begin{aligned}
A &= \frac{2T_F\alpha_s(\mu)}{3\pi v}\, [g_s(\mu)]^2 \sum_{n=0}^{\infty} A^{(n)}\left(\frac{\alpha_s(\mu)}{4\pi}\right)^n, \\
B &= \frac{2T_F\alpha_s(\mu)}{3\pi v}\, [g_s(\mu)]^2 \sum_{n=0}^{\infty} B^{(n)}\left(\frac{\alpha_s(\mu)}{4\pi}\right)^n.
\end{aligned}
$$

$$(B.65)$$

For simplicity, we only give explicit results for the tree-level amplitudes in this appendix. To reduce the length of expressions, we use the kinematic variables s_{ijk} defined by

$$s_{ijk} = (p_i + p_j + p_k)^2 = s_{ij} + s_{ik} + s_{jk}.$$

$$(B.66)$$

The $H + 2$ jets process is nonplanar, which means that we cannot remove all the relative phases in the amplitudes. It is therefore most convenient to keep all expressions in spinor helicity notation. We will explicitly demonstrate an example of the phases which appear in Eqs. (B.71) and (B.72).

438

$q\bar{q}\,q'\bar{q}'H$ **and** $q\bar{q}\,q\bar{q}H$

The tree-level amplitudes entering the Wilson coefficients $\vec{C}_{(+;\pm)}$ and $\vec{C}_{(+\pm)}$ in Eqs. (5.121) and (5.122) are

$$A^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-; 5_H) = -B^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-; 5_H) = \frac{1}{2}\left[\frac{\langle 24\rangle^2}{\langle 12\rangle\langle 34\rangle} + \frac{[13]^2}{[12][34]}\right],$$

$$A^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+; 5_H) = -B^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+; 5_H) = -\frac{1}{2}\left[\frac{\langle 23\rangle^2}{\langle 12\rangle\langle 34\rangle} + \frac{[14]^2}{[12][34]}\right], \quad \text{(B.67)}$$

$gg\,q\bar{q}H$

The tree-level amplitudes entering the Wilson coefficients $\vec{C}_{+-(+)}$, $\vec{C}_{++(+)}$, and $\vec{C}_{--(+)}$ in Eq. (5.126) are

$$A^{(0)}(1^+, 2^-; 3_q^+, 4_{\bar{q}}^-; 5_H) = \frac{\langle 24\rangle^3}{\langle 12\rangle\langle 14\rangle\langle 34\rangle} - \frac{[13]^3}{[12][23][34]},$$

$$A^{(0)}(2^-, 1^+; 3_q^+, 4_{\bar{q}}^-; 5_H) = \frac{[13]^2[14]}{[12][24][34]} - \frac{\langle 23\rangle\langle 24\rangle^2}{\langle 12\rangle\langle 13\rangle\langle 34\rangle},$$

$$A^{(0)}(1^+, 2^+; 3_q^+, 4_{\bar{q}}^-; 5_H) = -\frac{[1|2+3|4\rangle^2[23]}{s_{234}\langle 24\rangle}\left(\frac{1}{s_{23}} + \frac{1}{s_{34}}\right) + \frac{[2|1+3|4\rangle^2[13]}{s_{134}s_{34}\langle 14\rangle} - \frac{[3|1+2|4\rangle^2}{\langle 12\rangle\langle 14\rangle\langle 24\rangle[34]},$$

$$A^{(0)}(1^-, 2^-; 3_q^+, 4_{\bar{q}}^-; 5_H) = \frac{\langle 2|1+4|3]^2\langle 14\rangle}{s_{134}[13]}\left(\frac{1}{s_{14}} + \frac{1}{s_{34}}\right) - \frac{\langle 1|2+4|3]^2\langle 24\rangle}{s_{234}s_{34}[23]} + \frac{\langle 4|2+1|3]^2}{[12][13][23]\langle 34\rangle}.$$

$$\text{(B.68)}$$

In these expressions we have eliminated the Higgs momentum, p_5, using momentum conservation, so that all momenta appearing in the above expressions are lightlike. We have also used an extended spinor-helicity sandwich, defined by $[i|j+k|l\rangle = [i|j|l\rangle + [i|k|l\rangle$ to simplify notation.

All the B amplitudes vanish at tree level,

$$B^{(0)}(1^+, 2^-; 3_q^+, 4_{\bar{q}}^-; 5_H) = B^{(0)}(1^+, 2^+; 3_q^+, 4_{\bar{q}}^-; 5_H) = B^{(0)}(1^-, 2^-; 3_q^+, 4_{\bar{q}}^-; 5_H) = 0. \quad \text{(B.69)}$$

The tree-level amplitudes entering the Wilson coefficients $\vec{C}_{++--}$, $\vec{C}_{+++-}$, and $\vec{C}_{++++}$ in Eq. (5.133) are

$$A^{(0)}(1^+,2^+,3^+,4^+;5_H) = \frac{-2M_H^4}{\langle 12\rangle\langle 23\rangle\langle 34\rangle\langle 41\rangle},$$

$$A^{(0)}(1^+,2^+,3^+,4^-;5_H) = 2\left[\frac{[1|2+3|4\rangle^2[23]^2}{s_{234}s_{23}s_{34}} + \frac{[2|1+3|4\rangle^2[13]^2}{s_{134}s_{14}s_{34}} + \frac{[3|1+2|4\rangle^2[12]^2}{s_{124}s_{12}s_{14}}\right.$$
$$\left. + \frac{[13]}{[41]\langle 12\rangle\langle 23\rangle[34]}\left(\frac{s_{12}[1|2+3|4\rangle}{\langle 34\rangle} + \frac{s_{23}[3|1+2|4\rangle}{\langle 41\rangle} + [13]s_{123}\right)\right],$$

$$A^{(0)}(1^+,2^+,3^-,4^-;5_H) = 2\left[\frac{[12]^4}{[12][23][34][41]} + \frac{\langle 34\rangle^4}{\langle 12\rangle\langle 23\rangle\langle 34\rangle\langle 41\rangle}\right],$$

$$A^{(0)}(1^+,4^-,2^+,3^-;5_H) = 2\left[\frac{[12]^4}{[13][14][23][24]} + \frac{\langle 34\rangle^4}{\langle 13\rangle\langle 14\rangle\langle 23\rangle\langle 24\rangle}\right]. \tag{B.70}$$

To illustrate the relative phases that appear in these amplitudes, we can rewrite the amplitude $A^{(0)}(1^+,2^+,3^-,4^-;5_H)$ in terms of the Lorentz invariants s_{ij}

$$A^{(0)}(1^+,2^+,3^-,4^-;5_H) = 2e^{i\Phi_{++--H}}\left[\frac{s_{12}^2}{\sqrt{|s_{12}s_{23}s_{34}s_{14}|}} + e^{i\varphi}\frac{s_{34}^2}{\sqrt{|s_{12}s_{23}s_{34}s_{14}|}}\right], \tag{B.71}$$

with

$$\varphi = -2t\arg\left\{\frac{i\sqrt{s_{23}}[-s_{12}s_{34} + s_{13}s_{24} + s_{14}s_{23} - i(\sqrt{\alpha} + 2\sqrt{s_{13}}\sqrt{s_{23}}s_{14})]}{-s_{12}s_{34}(\sqrt{s_{13}} - i\sqrt{s_{23}}) + (s_{13}s_{24} - s_{14}s_{23} + i\sqrt{\alpha})(\sqrt{s_{13}} + i\sqrt{s_{23}})}\right\},$$

$$\alpha = 16(\epsilon_{\mu\nu\rho\sigma}p_1^\mu p_2^\nu p_3^\rho p_4^\sigma)^2 = 4s_{13}s_{14}s_{23}s_{24} - (s_{12}s_{34} - s_{13}s_{24} - s_{14}s_{23})^2 \geq 0,$$

$$\mathbf{t} = \text{sgn}(\epsilon_{\mu\nu\rho\sigma}p_1^\mu p_2^\nu p_3^\rho p_4^\sigma). \tag{B.72}$$

The branch cut of the square root is given by the usual prescription, $\sqrt{s_{ij}} \equiv \sqrt{s_{ij}+i0} = i\sqrt{|s_{ij}|}$ if $s_{ij} < 0$. Our convention for the antisymmetric Levi-Civita tensor is $\epsilon_{0123} = -1$. For this process we can choose a frame where all but one of the momenta p_1 through p_4 lie in a plane (with p_5 determined by momentum conservation). The phase φ is needed to determine the momentum of the nonplanar momentum and the sign $\mathbf{t}$ resolves which side of

the plane this particle is on, which is not captured by the s_{ij} (because they are symmetric with respect to a reflection about the plane). We note the simplicity of the spinor-helicity expression as compared with the explicit expression for the phases.

B.3 Helicity Amplitudes for Vector Boson + Jets

In this appendix we give all required partial amplitudes for the vector boson + jets processes discussed in Sec. 5.6. For each of the amplitudes $A_{q,v,a}$, $B_{q,v,a}$ defined in Sec. 5.6, we split the amplitude into its IR-divergent and IR-finite parts,

$$X = X_{\text{div}} + X_{\text{fin}}, \tag{B.73}$$

where X stands for any of $A_{q,v,a}$ and $B_{q,v,a}$. For the logarithms we use the notation

$$L_{ij} = \ln\left(-\frac{s_{ij}}{\mu^2} - i0\right), \qquad L_{ij/kl} = L_{ij} - L_{kl} = \ln\left(-\frac{s_{ij}}{\mu^2} - i0\right) - \ln\left(-\frac{s_{kl}}{\mu^2} - i0\right). \tag{B.74}$$

B.3.1 $V + 0$ Jets

In this section we give the amplitudes $A_{q,v,a}$ for $V + 0$ jets. For each partonic channel, we expand the amplitudes as

$$X = \sum_{n=0}^{\infty} X^{(n)} \left(\frac{\alpha_s(\mu)}{4\pi}\right)^n. \tag{B.75}$$

where X stands for any of $A_{q,v,a}$. The tree-level and one-loop helicity amplitudes entering the matching coefficient in Eq. (5.146) are given by

$$A_q^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_\ell^+, 4_{\bar{\ell}}^-) = -2i \frac{[13]\langle 24\rangle}{s_{12}},$$

$$A_{q,\text{div}}^{(1)}(1_q^+, 2_{\bar{q}}^-; 3_\ell^+, 4_{\bar{\ell}}^-) = A_q^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_\ell^+, 4_{\bar{\ell}}^-)\, C_F \left[-\frac{2}{\epsilon^2} + \frac{1}{\epsilon}\left(2L_{12} - 3\right)\right],$$

$$A_{q,\text{fin}}^{(1)}(1_q^+, 2_{\bar{q}}^-; 3_\ell^+, 4_{\bar{\ell}}^-) = A_q^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_\ell^+, 4_{\bar{\ell}}^-)\, C_F \left[-L_{12}^2 + 3L_{12} - 8 + \frac{\pi^2}{6}\right],$$

$$A_v^{(0)} = A_v^{(1)} = A_a^{(0)} = A_a^{(1)} = 0. \tag{B.76}$$

441

B.3.2 $V + 1$ Jet

In this section we give the amplitudes $A_{q,v,a}$ for $V + 1$ jets. Each amplitude is expanded as

$$X = g_s(\mu) \sum_{n=0}^{\infty} X^{(n)} \left(\frac{\alpha_s(\mu)}{4\pi} \right)^n \tag{B.77}$$

where X stands for any of $A_{q,v,a}$. The tree-level and one-loop helicity amplitudes for $V + 1$ jets were calculated in Refs. [232, 233, 362, 235]. We use the results given in Ref. [235], which uses $T_F = 1$ and $g_s T^a / \sqrt{2}$ for the $q\bar{q}g$ coupling. We can thus convert to our conventions by replacing $T^a \to \sqrt{2} T^a$, and identifying $1/N = C_A - 2C_F$ and $N = C_A$. The one-loop amplitudes are given in the FDH scheme in Ref. [235], which we convert to the HV scheme using Eqs. (5.100) and (5.101).

The tree-level amplitudes entering the matching coefficient $\vec{C}_{x+(+;+)}$ in Eq. (5.154) is given by

$$A_q^{(0)}(1^+; 2_q^+, 3_{\bar{q}}^-; 4_\ell^+, 5_{\bar{\ell}}^-) = -2\sqrt{2} \frac{\langle 35 \rangle^2}{\langle 12 \rangle \langle 13 \rangle \langle 45 \rangle}$$

$$A_v^{(0)} = A_a^{(0)} = 0. \tag{B.78}$$

The divergent part of the one-loop helicity amplitude is given by

$$A_{q,\text{div}}^{(1)}(1^+; 2_q^+, 3_{\bar{q}}^-; 4_\ell^+, 5_{\bar{\ell}}^-) = A_q^{(0)}(1^+; 2_q^+, 3_{\bar{q}}^-; 4_\ell^+, 5_{\bar{\ell}}^-)$$

$$\times \left\{ -\frac{1}{\epsilon^2}(C_A + 2C_F) + \frac{1}{\epsilon} \left[C_A(L_{12} + L_{13} - L_{23}) + C_F(2L_{23} - 3) - \frac{t_0}{2} \right] \right\}. \tag{B.79}$$

The finite parts entering the matching coefficients at one loop are

$$A_{q,\text{fin}}^{(1)}(1^+; 2_q^+, 3_{\bar{q}}^-; 4_\ell^+, 5_{\bar{\ell}}^-)$$

$$= A_q^{(0)}(1^+, 2_q^+, 3_{\bar{q}}^-; 4_\ell^+, 5_{\bar{\ell}}^-)$$

$$\times \left\{ \frac{C_A}{2} \left(-L_{12}^2 - L_{13}^2 + 3L_{13} - 7 + \frac{\pi^2}{3} \right) + \left(C_F - \frac{C_A}{2} \right) \left(-L_{23}^2 + 3L_{45} - 8 + \frac{\pi^2}{6} \right) \right.$$

442

$$+ C_A \left[-\mathrm{Ls}_{-1}\left(\frac{s_{12}}{s_{45}}, \frac{s_{13}}{s_{45}}\right) + \frac{\langle 3|24|5\rangle}{\langle 35\rangle s_{45}} \mathrm{L}_0\left(\frac{s_{13}}{s_{45}}\right) - \frac{1}{2}\frac{\langle 3|24|5\rangle^2}{\langle 35\rangle^2 s_{45}^2} \mathrm{L}_1\left(\frac{s_{13}}{s_{45}}\right)\right]$$

$$+ (C_A - 2C_F) \left[\mathrm{Ls}_{-1}\left(\frac{s_{12}}{s_{45}}, \frac{s_{23}}{s_{45}}\right) + \left(\frac{\langle 25\rangle^2\langle 13\rangle^2}{\langle 12\rangle^2\langle 35\rangle^2} - \frac{\langle 15\rangle\langle 23\rangle + \langle 13\rangle\langle 25\rangle}{\langle 12\rangle\langle 35\rangle}\right)\mathrm{Ls}_{-1}\left(\frac{s_{13}}{s_{45}}, \frac{s_{23}}{s_{45}}\right)\right.$$

$$+ \frac{2[12]\langle 25\rangle\langle 13\rangle}{\langle 35\rangle s_{45}}\mathrm{L}_0\left(\frac{s_{13}}{s_{45}}\right) + \frac{\langle 25\rangle^2[2|1|3\rangle\langle 13\rangle}{\langle 12\rangle\langle 35\rangle^2 s_{13}}\mathrm{L}_0\left(\frac{s_{45}}{s_{13}}\right) - \frac{\langle 3|21|5\rangle\langle 25\rangle\langle 13\rangle}{\langle 12\rangle\langle 35\rangle^2 s_{23}}\mathrm{L}_0\left(\frac{s_{45}}{s_{23}}\right)$$

$$- \frac{\langle 13\rangle^2[1|2|5\rangle^2}{2\langle 35\rangle^2 s_{13}^2}\mathrm{L}_1\left(\frac{s_{45}}{s_{13}}\right) - \frac{[1|2|3\rangle[4|1|5\rangle\langle 45\rangle\langle 13\rangle}{\langle 35\rangle^2 s_{23}^2}\mathrm{L}_1\left(\frac{s_{45}}{s_{23}}\right)$$

$$\left.\left. + \frac{[14]([12][34] + [13][24])\langle 13\rangle\langle 45\rangle}{2\langle 35\rangle^2[13][23][45]}\right]\right\},$$

$$A_a^{(1)}(1^+; 2_q^+, 3_{\bar{q}}^-; 4_\ell^+, 5_{\bar{\ell}}^-) = 4\sqrt{2} T_F [12][14]\langle 35\rangle \left[\frac{1}{s_{45}^2}\mathrm{L}_1\left(\frac{s_{23}}{s_{45}}\right) - \frac{1}{12 s_{45} m_t^2}\right],$$

$$A_v^{(1)} = 0. \tag{B.80}$$

The contributions from virtual top quark loops are calculated in an expansion in $1/m_t$ to order $1/m_t^2$ in Ref. [235], hence the divergent behavior of $A_a^{(1)}$ as $m_t \to 0$. To reduce the length of the expressions, we have used the commonly defined functions

$$\mathrm{L}_0(r) = \frac{\ln r}{1-r}, \qquad \mathrm{L}_1(r) = \frac{\mathrm{L}_0(r) + 1}{1-r}, \qquad \mathrm{Ls}_{-1}(r_1, r_2) = \mathrm{Li}_2(1 - r_1) + \mathrm{Li}_2(1 - r_2) + \ln r_1 \ln r_2 - \frac{\pi^2}{6}. \tag{B.81}$$

The proper branch cut of logarithms follows from the prescriptions $s_{ij} \to s_{ij} + \mathrm{i}0$. The proper branch cut of the dilogarithm follows from that of the logarithm through the identity

$$\mathrm{Im}\left[\mathrm{Li}_2(1-r)\right] = -\ln(1-r)\,\mathrm{Im}\left[\ln r\right]. \tag{B.82}$$

B.3.3 $V + 2$ Jets

In this section, we give the amplitudes $A_{q,v,a}$, $B_{q,v,a}$ for $V+2$ jets. Each amplitude is expanded as

$$X = [g_s(\mu)]^2 \sum_{n=0}^{\infty} X^{(n)}\left(\frac{\alpha_s(\mu)}{4\pi}\right)^n \tag{B.83}$$

443

where X stands for any of $A_{q,v,a}$ or $B_{q,v,a}$. We also define the kinematic variables s_{ijk} as

$$s_{ijk} = (p_i + p_j + p_k)^2 = s_{ij} + s_{ik} + s_{jk} \,. \qquad (B.84)$$

The one-loop helicity amplitudes for $q'\bar{q}'q\bar{q}V$ and $q\bar{q}q\bar{q}V$ were calculated in Ref. [367]. The one-loop helicity amplitudes for $gg\,q\bar{q}V$ were calculated in Ref. [235], which also gives compact expressions for the four-quark amplitudes, which we use here. The contributions from virtual top quark loops are calculated in an expansion in $1/m_t$ to order $1/m_t^2$ in Ref. [235].

Reference [235] uses $T_F = 1$ and $g_s T^a / \sqrt{2}$ for the $q\bar{q}g$ coupling. We can thus convert to our conventions by replacing $T^a \to \sqrt{2}T^a$, and identifying $1/N = C_A - 2C_F$ and $N = C_A$. The one-loop amplitudes are given in the FDH scheme in Ref. [235], which we convert to the HV scheme using Eqs. (5.100) and (5.101).

$q'\bar{q}'q\bar{q}V$ **and** $q\bar{q}q\bar{q}V$

The tree-level amplitudes for $q'\bar{q}'q\bar{q}V$ and $q\bar{q}q\bar{q}V$ entering the Wilson coefficients in Eqs. (5.173) and (5.174) are given by

$$A_q^{(0)}(1_{q'}^+, 2_{\bar{q}'}^-; 3_q^+, 4_{\bar{q}}^-; 5_\ell^+, 6_{\bar{\ell}}^-) = -B_q^{(0)}(1_{q'}^+, 2_{\bar{q}'}^-; 3_q^+, 4_{\bar{q}}^-; 5_\ell^+, 6_{\bar{\ell}}^-)$$

$$= \frac{2}{s_{12}s_{56}} \left[\frac{[13]\langle 46\rangle(\langle 12\rangle[15] - \langle 23\rangle[35])}{s_{123}} + \frac{\langle 24\rangle[35]([12]\langle 26\rangle + [14]\langle 46\rangle)}{s_{124}} \right],$$

$$A_q^{(0)}(1_{q'}^-, 2_{\bar{q}'}^+; 3_q^+, 4_{\bar{q}}^-; 5_\ell^+, 6_{\bar{\ell}}^-) = -B_q^{(0)}(1_{q'}^-, 2_{\bar{q}'}^+; 3_q^+, 4_{\bar{q}}^-; 5_\ell^+, 6_{\bar{\ell}}^-)$$

$$= \frac{2}{s_{12}s_{56}} \left[\frac{[23]\langle 46\rangle(\langle 12\rangle[25] + \langle 13\rangle[35])}{s_{123}} + \frac{\langle 14\rangle[35]([12]\langle 16\rangle - [24]\langle 46\rangle)}{s_{124}} \right],$$

$$A_v^{(0)} = A_a^{(0)} = B_v^{(0)} = B_a^{(0)} = 0 \,. \qquad (B.85)$$

Due to the length of the one-loop $q'\bar{q}'q\bar{q}V$ amplitudes, we only show how to translate the decomposition of the amplitude in Ref. [235] to our notation. The one-loop amplitudes are given in terms of the bare partial amplitudes $A_{i;j}(3_q, 2_{\bar{Q}}, 1_Q, 4_{\bar{q}})$ of Ref. [235] as

$$A_q^{(1)}(1_{q'}, 2_{\bar{q}'}; 3_q, 4_{\bar{q}}; 6_\ell^+, 5_{\bar{\ell}}^-) = -\mathrm{i}\,32\pi^2\,N\,A_{6;1}(3_q, 2_{\bar{Q}}, 1_Q, 4_{\bar{q}}) - \left(\frac{\beta_0}{\epsilon} + 2C_F - \frac{1}{3}C_A \right) A_q^{(0)}(1_{q'}, 2_{\bar{q}'}; 3_q, 4_{\bar{q}}; 6_\ell^+, 5_{\bar{\ell}}^-),$$

$$B_q^{(1)}(1_{q'}, 2_{\bar{q}'}; 3_q, 4_{\bar{q}}; 6_\ell^+, 5_{\bar{\ell}}^-) = -\mathrm{i}\,32\pi^2\, N\, A_{6;2}(3_q, 2_{\bar{Q}}, 1_Q, 4_{\bar{q}}) - \left(\frac{\beta_0}{\epsilon} + 2C_F - \frac{1}{3}C_A\right) B_q^{(0)}(1_{q'}, 2_{\bar{q}'}; 3_q, 4_{\bar{q}}; 6_\ell^+, 5_{\bar{\ell}}^-)\,,$$

$$A_a^{(1)}(1_{q'}, 2_{\bar{q}'}; 3_q, 4_{\bar{q}}; 6_\ell^+, 5_{\bar{\ell}}^-) = -B_a^{(1)}(1_{q'}, 2_{\bar{q}'}; 3_q, 4_{\bar{q}}; 6_\ell^+, 5_{\bar{\ell}}^-) = -\mathrm{i}\,32\pi^2\, A_{6;3}(3_q, 2_{\bar{Q}}, 1_Q, 4_{\bar{q}})\,,$$

$$A_v^{(1)} = B_v^{(1)} = 0\,. \tag{B.86}$$

The overall factor $-\mathrm{i}\,32\pi^2$ is due to our different normalization conventions. We have not included helicity labels, as these relations are true for all helicity combinations. Note that the partial amplitudes $A_{i;j}$ do not include labels for the lepton momenta, which are implicitly taken as 6_ℓ^+, $5_{\bar{\ell}}^-$. The terms in the first two lines proportional to $A_q^{(0)}$ and $B_q^{(0)}$ come from the UV renormalization and switching from FDH to HV.

$gg\,q\bar{q}\,V$

The tree-level amplitudes for $gg\,q\bar{q}\,V$ entering the matching coefficients in Eq. (5.182) are given by

$$A_q^{(0)}(1^+, 2^+; 3_q^+, 4_{\bar{q}}^-; 5_\ell^+, 6_{\bar{\ell}}^-) = -4\frac{\langle 46\rangle^2}{\langle 12\rangle\langle 13\rangle\langle 24\rangle\langle 56\rangle}\,,$$

$$A_q^{(0)}(1^+, 2^-; 3_q^+, 4_{\bar{q}}^-; 5_\ell^+, 6_{\bar{\ell}}^-) = \frac{4}{s_{12}s_{56}}\left[\frac{[13]\langle 23\rangle\langle 46\rangle(\langle 23\rangle[35] - \langle 12\rangle[15])}{\langle 13\rangle s_{123}} + \frac{\langle 24\rangle[35][14]([12]\langle 26\rangle + [14]\langle 46\rangle)}{[24]s_{124}}\right.$$
$$\left. + \frac{([12]\langle 26\rangle + [14]\langle 46\rangle)(\langle 23\rangle[35] - \langle 12\rangle[15])}{\langle 13\rangle[24]}\right]\,,$$

$$A_q^{(0)}(1^-, 2^+; 3_q^+, 4_{\bar{q}}^-; 5_\ell^+, 6_{\bar{\ell}}^-) = \frac{4}{s_{12}s_{56}}\left[\frac{[23]^2\langle 46\rangle(\langle 12\rangle[25] + \langle 13\rangle[35])}{[13]s_{123}} + \frac{\langle 14\rangle^2[35]([12]\langle 16\rangle - [24]\langle 46\rangle)}{\langle 24\rangle s_{124}}\right.$$
$$\left. + \frac{[23]\langle 14\rangle[35]\langle 46\rangle}{[13]\langle 24\rangle}\right]\,,$$

$$A_v^{(0)} = A_a^{(0)} = B_q^{(0)} = B_v^{(0)} = B_a^{(0)} = 0\,. \tag{B.87}$$

Due to the length of the one-loop $gg\,q\bar{q}\,V$ amplitudes, we again only show how to translate the decomposition of the amplitude in Ref. [235] to our notation. The one-loop amplitudes are given in terms of the bare partial amplitudes $A_{i;j}(3_q, 1, 2, 4_{\bar{q}})$, $A_{i;j}^v(3_q, 4_{\bar{q}}; 1, 2)$, and

$A^{ax}_{i;j}(3_q, 4_{\bar{q}}; 1, 2)$ of Ref. [235] as

$$A^{(1)}_q(1, 2; 3_q, 4_{\bar{q}}; 6^+_\ell, 5^-_{\bar{\ell}}) = -i\,64\pi^2\, N\, A_{6;1}(3_q, 1, 2, 4_{\bar{q}}) - \left(\frac{\beta_0}{\epsilon} + C_F\right) A^{(0)}_q(1, 2; 3_q, 4_{\bar{q}}; 6^+_\ell, 5^-_{\bar{\ell}}),$$

$$B^{(1)}_q(1, 2; 3_q, 4_{\bar{q}}; 6^+_\ell, 5^-_{\bar{\ell}}) = -i\,64\pi^2\, A_{6;3}(3_q, 4_{\bar{q}}; 1, 2),$$

$$A^{(1)}_v(1, 2; 3_q, 4_{\bar{q}}; 6^+_\ell, 5^-_{\bar{\ell}}) = -i\,64\pi^2\, A^v_{6;4}(3_q, 4_{\bar{q}}; 1, 2),$$

$$B^{(1)}_v(1, 2; 3_q, 4_{\bar{q}}; 6^+_\ell, 5^-_{\bar{\ell}}) = +i\,64\pi^2\, \frac{2}{N}\, A^v_{6;4}(3_q, 4_{\bar{q}}; 1, 2),$$

$$A^{(1)}_a(1, 2; 3_q, 4_{\bar{q}}; 6^+_\ell, 5^-_{\bar{\ell}}) = -i\,64\pi^2\, A^{ax}_{6;4}(3_q, 4_{\bar{q}}; 1, 2),$$

$$B^{(1)}_a(1, 2; 3_q, 4_{\bar{q}}; 6^+_\ell, 5^-_{\bar{\ell}}) = -i\,64\pi^2\, \frac{1}{N}\, \left[A^{ax}_{6;5}(3_q, 4_{\bar{q}}; 1, 2) - A^{ax}_{6;4}(3_q, 4_{\bar{q}}; 1, 2) - A^{ax}_{6;4}(3_q, 4_{\bar{q}}; 2, 1)\right].$$

$$(\text{B.88})$$

The overall factor $-i\,64\pi^2$ is due to our different normalization conventions. We have not included helicity labels, as these relations are true for all helicity combinations. Note that the partial amplitudes $A_{i;j}$ do not include labels for the lepton momenta, which are implicitly taken as 6^+_ℓ, $5^-_{\bar{\ell}}$. The term in the first line proportional to $A^{(0)}_q$ comes from the UV renormalization and switching from FDH to HV.

B.4 Helicity Amplitudes for $pp \to$ Jets

B.4.1 $pp \to 2$ Jets

In this appendix we give explicit expressions for all partial amplitudes that are required in Eqs. (5.189), (5.190), (5.194), and (5.203), for the various partonic channels of the $pp \to 2$ jets process. Since this process is planar, we can write all amplitudes for a given set of helicities with a common overall phase extracted, which is determined by the phases of the external particles. In this way, we do not need to worry about relative phases between the Wilson coefficients for different color structures when they mix under renormalization. The cross section does not depend on this overall phase. This simplifies the numerical implementation considerably for this process, as it avoids having to implement the complex spinor algebra.

To extract the overall phase from the amplitudes, the following relations for the relative phases between the spinor products are useful:

$$\frac{\langle 12 \rangle}{[34]} = \frac{\langle 34 \rangle}{[12]} = \frac{\langle 14 \rangle}{[23]} = \frac{\langle 23 \rangle}{[14]} = -\frac{\langle 13 \rangle}{[24]} = -\frac{\langle 24 \rangle}{[13]} \,. \tag{B.89}$$

These relations follow from Eq. (B.22) with $n = 4$.

We split the partial amplitudes into their IR-divergent and IR-finite parts,

$$A = A_{\mathrm{div}} + A_{\mathrm{fin}} \,, \qquad B = B_{\mathrm{div}} + B_{\mathrm{fin}} \,, \tag{B.90}$$

where the IR-finite parts enter the matching coefficients. We expand the amplitudes and Wilson coefficients in powers of $\alpha_s(\mu)$ as

$$X = [g_s(\mu)]^2 \sum_{n=0}^{\infty} X^{(n)} \left(\frac{\alpha_s(\mu)}{4\pi} \right)^n \,, \tag{B.91}$$

where X stands for any of $A_{\mathrm{div,fin}}$, $B_{\mathrm{div,fin}}$, and $X^{(0)}$ and $X^{(1)}$ are the tree-level and one-loop contributions, respectively. For simplicity, we drop the subscript "fin" for those amplitudes that have no divergent parts, e.g., for the tree-level amplitudes $A_{\mathrm{div}}^{(0)} = 0$ and $A_{\mathrm{fin}}^{(0)} \equiv A^{(0)}$. For the logarithms we use the notation

$$L_{ij} = \ln\left(-\frac{s_{ij}}{\mu^2} - i0 \right) , \qquad L_{ij/kl} = L_{ij} - L_{kl} = \ln\left(-\frac{s_{ij}}{\mu^2} - i0 \right) - \ln\left(-\frac{s_{kl}}{\mu^2} - i0 \right) . \tag{B.92}$$

$q\bar{q}\,q'\bar{q}'$ and $q\bar{q}\,q\bar{q}$

Here we list all partial amplitudes up to one loop entering the Wilson coefficients in Eqs. (5.189) and (5.190). The one-loop helicity amplitudes for $q\bar{q}\,q'\bar{q}'$ and $q\bar{q}\,q\bar{q}$ were first calculated in Ref. [244], and the two-loop helicity amplitudes were computed in Refs. [372, 373]. We find agreement between the one-loop results of Refs. [372] and [373], from which we take our results.[1] Our one-loop matching coefficients agree with the calculation of Ref. [296].

---

[1]Note that there is a minor disagreement here with the earlier calculation in Ref. [244], presumably due to typos. Specifically, in Ref. [244] the factors $\left(\log^2 \frac{s_{14}}{s_{12}} + \pi^2 \right)$ and $\left(\log^2 \frac{s_{14}}{s_{13}} + \pi^2 \right)$ in Eqs. (5.10) and (5.12)

To extract the overall phase from the amplitudes, the following relations for the relative phases between the spinor products are useful:

$$\frac{\langle 12 \rangle}{[34]} = \frac{\langle 34 \rangle}{[12]} = \frac{\langle 14 \rangle}{[23]} = \frac{\langle 23 \rangle}{[14]} = -\frac{\langle 13 \rangle}{[24]} = -\frac{\langle 24 \rangle}{[13]} \,. \tag{B.89}$$

These relations follow from Eq. (B.22) with $n = 4$.

We split the partial amplitudes into their IR-divergent and IR-finite parts,

$$A = A_{\mathrm{div}} + A_{\mathrm{fin}} \,, \qquad B = B_{\mathrm{div}} + B_{\mathrm{fin}} \,, \tag{B.90}$$

where the IR-finite parts enter the matching coefficients. We expand the amplitudes and Wilson coefficients in powers of $\alpha_s(\mu)$ as

$$X = [g_s(\mu)]^2 \sum_{n=0}^{\infty} X^{(n)} \left(\frac{\alpha_s(\mu)}{4\pi} \right)^n \,, \tag{B.91}$$

where X stands for any of $A_{\mathrm{div,fin}}$, $B_{\mathrm{div,fin}}$, and $X^{(0)}$ and $X^{(1)}$ are the tree-level and one-loop contributions, respectively. For simplicity, we drop the subscript "fin" for those amplitudes that have no divergent parts, e.g., for the tree-level amplitudes $A_{\mathrm{div}}^{(0)} = 0$ and $A_{\mathrm{fin}}^{(0)} \equiv A^{(0)}$. For the logarithms we use the notation

$$L_{ij} = \ln\left(-\frac{s_{ij}}{\mu^2} - i0 \right) , \qquad L_{ij/kl} = L_{ij} - L_{kl} = \ln\left(-\frac{s_{ij}}{\mu^2} - i0 \right) - \ln\left(-\frac{s_{kl}}{\mu^2} - i0 \right) . \tag{B.92}$$

$q\bar{q}\,q'\bar{q}'$ and $q\bar{q}\,q\bar{q}$

Here we list all partial amplitudes up to one loop entering the Wilson coefficients in Eqs. (5.189) and (5.190). The one-loop helicity amplitudes for $q\bar{q}\,q'\bar{q}'$ and $q\bar{q}\,q\bar{q}$ were first calculated in Ref. [244], and the two-loop helicity amplitudes were computed in Refs. [372, 373]. We find agreement between the one-loop results of Refs. [372] and [373], from which we take our results.[1] Our one-loop matching coefficients agree with the calculation of Ref. [296].

---

[1]Note that there is a minor disagreement here with the earlier calculation in Ref. [244], presumably due to typos. Specifically, in Ref. [244] the factors $\left(\log^2 \frac{s_{14}}{s_{12}} + \pi^2 \right)$ and $\left(\log^2 \frac{s_{14}}{s_{13}} + \pi^2 \right)$ in Eqs. (5.10) and (5.12)

The tree-level amplitudes are

$$A^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-) = -B^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-) = -\frac{\langle 24\rangle[13]}{s_{12}} = \frac{s_{13}}{s_{12}} e^{i\Phi_{(+;+)}},$$

$$A^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+) = -B^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+) = -\frac{\langle 23\rangle[14]}{s_{12}} = \frac{s_{14}}{s_{12}} e^{i\Phi_{(+;-)}}, \qquad (\text{B.93})$$

where the phases are given by

$$e^{i\Phi_{(+;+)}} = \frac{\langle 24\rangle}{\langle 13\rangle}, \qquad e^{i\Phi_{(+;-)}} = \frac{\langle 23\rangle}{\langle 14\rangle}. \qquad (\text{B.94})$$

We have chosen to express all the one-loop amplitudes in terms of $A^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-)$ and $A^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+)$. The divergent parts of the one-loop amplitudes are

$$A_{\text{div}}^{(1)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-) = A^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-)\left\{ -\frac{4}{\epsilon^2}C_F + \frac{2}{\epsilon}\left[C_F(2L_{12} - 4L_{13/14} - 3) + C_A(L_{13/14} - L_{12/13}) \right] \right\},$$

$$B_{\text{div}}^{(1)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-) = A^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-)\left\{ \frac{4}{\epsilon^2}C_F - \frac{2}{\epsilon}\left[C_F(2L_{12} - 2L_{13/14} - 3) + C_A(L_{13/14} - L_{12/13}) \right] \right\},$$

$$A_{\text{div}}^{(1)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+) = A^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+)\left\{ -\frac{4}{\epsilon^2}C_F + \frac{2}{\epsilon}\left[C_F(2L_{12} - 4L_{13/14} - 3) + C_A(L_{13/14} - L_{12/13}) \right] \right\},$$

$$B_{\text{div}}^{(1)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+) = A^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+)\left\{ \frac{4}{\epsilon^2}C_F - \frac{2}{\epsilon}\left[C_F(2L_{12} - 2L_{13/14} - 3) + C_A(L_{13/14} - L_{12/13}) \right] \right\}.$$

$$(\text{B.95})$$

The finite parts entering the Wilson coefficients are

$$A_{\text{fin}}^{(1)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-) = A^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-)\left[f(s_{12}, s_{13}, s_{14}, \mu) + (4C_F - C_A)\, g(s_{12}, s_{13}, s_{14}) \right],$$

$$B_{\text{fin}}^{(1)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-) = A^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^+, 4_{\bar{q}'}^-)\left[4C_F\, L_{12}L_{13/14} - f(s_{12}, s_{13}, s_{14}, \mu) + (C_A - 2C_F)\, g(s_{12}, s_{13}, s_{14}) \right],$$

$$A_{\text{fin}}^{(1)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+) = A^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+)\left[f(s_{12}, s_{13}, s_{14}, \mu) + 2(C_A - 2C_F)\, g(s_{12}, s_{14}, s_{13}) \right],$$

$$B_{\text{fin}}^{(1)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+) = A^{(0)}(1_q^+, 2_{\bar{q}}^-; 3_{q'}^-, 4_{\bar{q}'}^+)\left[4C_F\, L_{12}L_{13/14} - f(s_{12}, s_{13}, s_{14}, \mu) + 2(C_F - C_A)\, g(s_{12}, s_{14}, s_{13}) \right],$$

$$f(s_{12}, s_{13}, s_{14}, \mu) = C_F\left[-2L_{12}^2 + 2L_{12}(3 + 4L_{13/14}) - 16 + \frac{\pi^2}{3} \right] + C_A\left(2L_{12}(L_{12/13} - L_{13/14}) + \frac{10}{3} + \pi^2 \right)$$

respectively, must be swapped to achieve agreement with the results of Refs. [372, 373, 296]. Reference [373] also has a minor typo, having a flipped overall sign for the IR-divergent terms.

448

$$- \beta_0 \left(L_{12} - \frac{5}{3} \right) ,$$

$$g(s_{12}, s_{13}, s_{14}) = \frac{s_{12}}{s_{13}} \left[\frac{1}{2} \left(1 - \frac{s_{14}}{s_{13}} \right) \left(L_{12/14}^2 + \pi^2 \right) + L_{12/14} \right] . \tag{B.96}$$

$gg\,q\bar{q}$

The one-loop helicity amplitudes for $gg\,q\bar{q}$ were first calculated in Ref. [244], and the two-loop helicity amplitudes were computed in Refs. [368, 369]. We take our results from Ref. [368], converted to our conventions.[2]

Here we list all partial amplitudes up to one loop entering the Wilson coefficients in Eq. (5.194). We start with the partial amplitudes where the gluons have opposite helicity, which are the only ones having a nonzero tree-level contribution. The tree-level amplitudes are given by

$$A^{(0)}(1^+, 2^-; 3_q^+, 4_{\bar{q}}^-) = -2 \frac{\langle 23 \rangle \langle 24 \rangle^3}{\langle 12 \rangle \langle 24 \rangle \langle 43 \rangle \langle 31 \rangle} = 2 \frac{\sqrt{|s_{13}\,s_{14}|}}{s_{12}} e^{i\Phi_{+-(+)}} ,$$

$$A^{(0)}(2^-, 1^+; 3_q^+, 4_{\bar{q}}^-) = -2 \frac{\langle 23 \rangle \langle 24 \rangle^3}{\langle 21 \rangle \langle 14 \rangle \langle 43 \rangle \langle 32 \rangle} = 2 \frac{s_{13}\sqrt{|s_{13}\,s_{14}|}}{s_{12}\,s_{14}} e^{i\Phi_{+-(+)}} ,$$

$$B^{(0)}(1^+, 2^-; 3_q^+, 4_{\bar{q}}^-) = 0 . \tag{B.97}$$

In the second step we extracted a common overall phase from the amplitudes, which is given by

$$e^{i\Phi_{+-(+)}} = \frac{\langle 24 \rangle}{[24]} \frac{[13][14]}{\sqrt{|s_{13}\,s_{14}|}} . \tag{B.98}$$

The divergent parts of the corresponding one-loop amplitudes are

$$A_{\mathrm{div}}^{(1)}(1^+, 2^-; 3_q^+, 4_{\bar{q}}^-) = A^{(0)}(1^+, 2^-; 3_q^+, 4_{\bar{q}}^-) \left[-\frac{2}{\epsilon^2} \left(C_A + C_F \right) + \frac{1}{\epsilon} \left(2C_F\,L_{12} + 2C_A\,L_{13} - 3C_F - \beta_0 \right) \right] ,$$

$$A_{\mathrm{div}}^{(1)}(2^-, 1^+; 3_q^+, 4_{\bar{q}}^-) = A^{(0)}(2^-, 1^+; 3_q^+, 4_{\bar{q}}^-) \left[-\frac{2}{\epsilon^2} \left(C_A + C_F \right) + \frac{1}{\epsilon} \left(2C_F\,L_{12} + 2C_A\,L_{14} - 3C_F - \beta_0 \right) \right] ,$$

---

[2]We find a slight disagreement with the earlier results of Ref. [244] for their subleading color amplitude in Eq. (5.24). This amplitude appears to have typos since it does not have the correct IR structure, as determined by the general formula [492] or by the SCET result in Eq. (B.148). Comparing with the matching calculation of Ref. [296], we find a typo in the π^2 term in W_4 in Eq. (54), which should have $3\pi^2 u^2/(2ts) \to -3\pi^2 u/(4t)$.

$$B_{\text{div}}^{(1)}(1^+, 2^-; 3_q^+, 4_{\bar q}^-) = A^{(0)}(1^+, 2^-; 3_q^+, 4_{\bar q}^-)\frac{1}{\epsilon}\, 4T_F\left(L_{12/14} + \frac{s_{13}}{s_{14}}L_{12/13}\right). \tag{B.99}$$

The corresponding finite parts entering the Wilson coefficient $\vec{C}_{+-(+)}$ at one loop are

$$A_{\text{fin}}^{(1)}(1^+, 2^-; 3_q^+, 4_{\bar q}^-) = A^{(0)}(1^+, 2^-; 3_q^+, 4_{\bar q}^-)\Bigg\{ C_A\left(-L_{13}^2 + L_{12/13}^2 + 1 + \frac{7\pi^2}{6}\right) + C_F\left(-L_{12}^2 + 3L_{12} - 8 + \frac{\pi^2}{6}\right)$$

$$+ (C_A - C_F)\frac{s_{12}}{s_{14}}\left(L_{12/13}^2 + \pi^2\right)\Bigg\},$$

$$A_{\text{fin}}^{(1)}(2^-, 1^+; 3_q^+, 4_{\bar q}^-) = A^{(0)}(2^-, 1^+; 3_q^+, 4_{\bar q}^-)\Bigg\{ \frac{C_A}{2}\left(-2L_{14}^2 + L_{12/14}^2 - 3\,L_{12/14} + 1 + \frac{4\pi^2}{3}\right)$$

$$+ C_F\left(-L_{12}^2 + 3L_{12} - 8 + \frac{\pi^2}{6}\right) - \frac{C_A}{2}\frac{s_{14}}{s_{13}}\left[\left(1 - \frac{s_{14}}{s_{13}}L_{12/14}\right)^2 + L_{12/14} + \frac{s_{14}^2}{s_{13}^2}\pi^2\right]$$

$$+ \left(\frac{C_A}{2} - C_F\right)\frac{s_{12}}{s_{13}}\left[\left(1 + \frac{s_{12}}{s_{13}}L_{12/14}\right)^2 - L_{12/14} + \frac{s_{12}^2}{s_{13}^2}\pi^2\right]\Bigg\},$$

$$B_{\text{fin}}^{(1)}(1^+, 2^-; 3_q^+, 4_{\bar q}^-) = A^{(0)}(1^+, 2^-; 3_q^+, 4_{\bar q}^-)\,4T_F\left[-L_{12}L_{13/14} + \frac{s_{12}}{s_{14}}L_{14}L_{12/13} - \frac{3}{4}\frac{s_{12}}{s_{13}}\left(L_{12/14}^2 + \pi^2\right)\right]. \tag{B.100}$$

The partial amplitudes where both gluons have the same helicity vanish at tree level,

$$A^{(0)}(1^+, 2^+; 3_q^+, 4_{\bar q}^-) = B^{(0)}(1^+, 2^+; 3_q^+, 4_{\bar q}^-) = 0,$$

$$A^{(0)}(1^-, 2^-; 3_q^+, 4_{\bar q}^-) = B^{(0)}(1^-, 2^-; 3_q^+, 4_{\bar q}^-) = 0. \tag{B.101}$$

The corresponding one-loop amplitudes entering the Wilson coefficients $\vec{C}_{++(+)}$ and $\vec{C}_{--(+)}$ are IR finite. They are

$$A^{(1)}(1^+, 2^+; 3_q^+, 4_{\bar q}^-) = 2\sqrt{|s_{13}\,s_{14}|}\,e^{i\Phi_{++(+)}}\left[(C_A - C_F)\frac{1}{s_{13}} + \frac{1}{3}(C_A - 2T_F\,n_f)\frac{1}{s_{12}}\right],$$

$$A^{(1)}(2^+, 1^+; 3_q^+, 4_{\bar q}^-) = -2\sqrt{|s_{13}\,s_{14}|}\,e^{i\Phi_{++(+)}}\left[(C_A - C_F)\frac{1}{s_{14}} + \frac{1}{3}(C_A - 2T_F\,n_f)\frac{1}{s_{12}}\right],$$

$$B^{(1)}(1^+, 2^+; 3_q^+, 4_{\bar q}^-) = 0, \tag{B.102}$$

and

$$A^{(1)}(1^-,2^-;3_q^+,4_{\bar{q}}^-) = 2\sqrt{|s_{13}\,s_{14}|}\,e^{i\Phi_{--(+)}}\left[(C_A - C_F)\frac{1}{s_{13}} + \frac{1}{3}(C_A - 2T_F\,n_f)\frac{1}{s_{12}}\right],$$

$$A^{(1)}(2^-,1^-;3_q^+,4_{\bar{q}}^-) = -2\sqrt{|s_{13}\,s_{14}|}\,e^{i\Phi_{--(+)}}\left[(C_A - C_F)\frac{1}{s_{14}} + \frac{1}{3}(C_A - 2T_F\,n_f)\frac{1}{s_{12}}\right],$$

$$B^{(1)}(1^-,2^-;3_q^+,4_{\bar{q}}^-) = 0,\qquad\qquad\qquad\qquad\qquad\qquad (\text{B}.103)$$

with the overall phases

$$e^{i\Phi_{++(+)}} = \frac{[12]}{\langle 12\rangle}\,\frac{[13]\langle 14\rangle}{\sqrt{|s_{13}\,s_{14}|}}\,,\qquad e^{i\Phi_{--(+)}} = \frac{\langle 12\rangle}{[12]}\,\frac{[13]\langle 14\rangle}{\sqrt{|s_{13}\,s_{14}|}}\,. \qquad (\text{B}.104)$$

gggg

The one-loop helicity amplitudes for *gggg* were first calculated in Ref. [244], and the two-loop amplitudes were computed in Refs. [374, 375]. The results given here are taken from Ref. [375], and converted to our conventions. We also find complete agreement with the expressions given in Ref. [244].[3]

The amplitudes inherit the cyclic symmetry of the traces, which means that many of the amplitudes appearing in Eq. (5.203) are related, for example

$$A(1^+,3^-,4^-,2^+) = A(2^+,1^+,3^-,4^-)\,. \qquad\qquad\qquad (\text{B}.105)$$

For the convenience of the reader, we will explicitly give all amplitudes needed in Eq. (5.203). We start with the partial amplitudes with two positive-helicity and two negative-helicity gluons, which are the only nonvanishing amplitudes at tree level. We have

$$A^{(0)}(1^+,2^+,3^-,4^-) = 4\frac{\langle 34\rangle^4}{\langle 12\rangle\langle 23\rangle\langle 34\rangle\langle 41\rangle} = 4\frac{s_{12}}{s_{14}}\,e^{i\Phi_{++--}}\,,$$

---

[3]We have also compared with the matching calculation of Ref. [296], which has a minor typo. In particular, in $\mathcal{F}(s,t,u)$ in Eq. (61) the n_f terms must be dropped and β_0 set to $11C_A/3$. Also as noted in Ref. [297], the last column of Table 5 in Ref. [296] applies to helicities 7, 8, while the second-to-last column applies to helicities 9–16.

$$A^{(0)}(1^+, 3^-, 4^-, 2^+) = 4\frac{\langle 34 \rangle^4}{\langle 13 \rangle \langle 34 \rangle \langle 42 \rangle \langle 21 \rangle} = 4\frac{s_{12}}{s_{13}} e^{i\Phi_{++--}} \,,$$

$$A^{(0)}(1^+, 4^-, 2^+, 3^-) = 4\frac{\langle 34 \rangle^4}{\langle 14 \rangle \langle 42 \rangle \langle 23 \rangle \langle 31 \rangle} = 4\frac{s_{12}^2}{s_{13}\, s_{14}} e^{i\Phi_{++--}} \,, \qquad (\text{B.106})$$

with the common overall phase

$$e^{i\Phi_{++--}} = -\frac{[12]}{\langle 12 \rangle}\frac{\langle 34 \rangle}{[34]} \,. \qquad (\text{B.107})$$

The corresponding $B^{(0)}$ all vanish,

$$B^{(0)}(1^+, 2^+, 3^-, 4^-) = B^{(0)}(1^+, 3^-, 4^-, 2^+) = B^{(0)}(1^+, 4^-, 2^+, 3^-) = 0 \,. \qquad (\text{B.108})$$

At one loop the $B^{(1)}$ amplitudes can be expressed in terms of the n_f-independent part of the $A^{(1)}$,

$$B^{(1)}(1^+, 2^+, 3^-, 4^-) = \frac{2T_F}{C_A} 2\left[A^{(1)}(1^+, 2^+, 3^-, 4^-) + A^{(1)}(1^+, 3^-, 4^-, 2^+) + A^{(1)}(1^+, 4^-, 2^+, 3^-) \right]\Big|_{n_f=0} \,.$$
$$(\text{B.109})$$

The same relation also holds for the other helicity assignments. Using the cyclic symmetries of the amplitudes, it follows that the last three entries in the Wilson coefficients in Eq. (5.203) at one loop are all equal to each other and are given by $2T_F/C_A$ times the sum of the first three entries at $n_f = 0$. The divergent parts of the one-loop amplitudes are

$$A^{(1)}_{\text{div}}(1^+, 2^+, 3^-, 4^-) = A^{(0)}(1^+, 2^+, 3^-, 4^-)\left[-\frac{4}{\epsilon^2} C_A + \frac{2}{\epsilon}(C_A L_{12} + C_A L_{14} - \beta_0) \right] ,$$

$$A^{(1)}_{\text{div}}(1^+, 3^-, 4^-, 2^+) = A^{(0)}(1^+, 3^-, 4^-, 2^+)\left[-\frac{4}{\epsilon^2} C_A + \frac{2}{\epsilon}(C_A L_{12} + C_A L_{13} - \beta_0) \right] ,$$

$$A^{(1)}_{\text{div}}(1^+, 4^-, 2^+, 3^-) = A^{(0)}(1^+, 4^-, 2^+, 3^-)\left[-\frac{4}{\epsilon^2} C_A + \frac{2}{\epsilon}(C_A L_{13} + C_A L_{14} - \beta_0) \right] ,$$

$$B^{(1)}_{\text{div}}(1^+, 2^+, 3^-, 4^-) = A^{(0)}(1^+, 2^+, 3^-, 4^-)\left[\frac{8T_F}{\epsilon}\left(L_{12/13} + \frac{s_{14}}{s_{13}} L_{12/14} \right) \right] ,$$

$$B^{(1)}_{\text{div}}(1^+, 3^-, 4^-, 2^+) = A^{(0)}(1^+, 2^+, 3^-, 4^-)\left[\frac{8T_F}{\epsilon}\left(\frac{s_{14}}{s_{13}} L_{13/14} + \frac{s_{12}}{s_{13}} L_{13/12} \right) \right] ,$$

$$B_{\text{div}}^{(1)}(1^+, 4^-, 2^+, 3^-) = A^{(0)}(1^+, 2^+, 3^-, 4^-)\left[\frac{8T_F}{\epsilon}\left(L_{14/13} + \frac{s_{12}}{s_{13}}L_{14/12}\right)\right].$$ (B.110)

The finite parts entering the Wilson coefficient $\vec{C}_{++--}$ at one loop are

$$A_{\text{fin}}^{(1)}(1^+, 2^+, 3^-, 4^-) = A^{(0)}(1^+, 2^+, 3^-, 4^-)\left[C_A\left(-2L_{12}L_{14} - \frac{4}{3} + \frac{4\pi^2}{3}\right) + \beta_0\left(L_{14} - \frac{5}{3}\right)\right],$$

$$A_{\text{fin}}^{(1)}(1^+, 3^-, 4^-, 2^+) = A^{(0)}(1^+, 3^-, 4^-, 2^+)\left[C_A\left(-2L_{12}L_{13} - \frac{4}{3} + \frac{4\pi^2}{3}\right) + \beta_0\left(L_{13} - \frac{5}{3}\right)\right],$$

$$A_{\text{fin}}^{(1)}(1^+, 4^-, 2^+, 3^-) = A^{(0)}(1^+, 4^-, 2^+, 3^-)\left\{C_A\left(-2L_{14}L_{13} + \frac{4}{3}\pi^2 - \frac{4}{3}\right) - \beta_0\left(\frac{5}{3} + \frac{s_{13}}{s_{12}}L_{14} + \frac{s_{14}}{s_{12}}L_{13}\right)\right.$$
$$- (C_A - 2T_F n_f)\frac{s_{13}s_{14}}{s_{12}^2}\left[1 + \left(\frac{s_{13}}{s_{12}} - \frac{s_{14}}{s_{12}}\right)L_{13/14} + \left(2 - \frac{s_{13}s_{14}}{s_{12}^2}\right)(L_{13/14}^2 + \pi^2)\right]$$
$$\left. - 3T_F n_f \frac{s_{13}s_{14}}{s_{12}^2}\left(L_{13/14}^2 + \pi^2\right)\right\},$$

$$B_{\text{fin}}^{(1)}(1^+, 2^+, 3^-, 4^-) = B_{\text{fin}}^{(1)}(1^+, 3^-, 4^-, 2^+) = B_{\text{fin}}^{(1)}(1^+, 4^-, 2^+, 3^-)$$

$$= -4T_F A^{(0)}(1^+, 2^+, 3^-, 4^-)\left[\frac{s_{14}}{s_{13}}2L_{13}L_{12/14} + 2L_{14}L_{12/13} + \frac{s_{14}}{s_{12}} + \frac{s_{14}}{s_{12}}\left(\frac{s_{13}}{s_{12}} - \frac{s_{14}}{s_{12}}\right)L_{13/14}\right.$$
$$\left. + \frac{s_{14}}{s_{12}}\left(2 - \frac{s_{13}s_{14}}{s_{12}^2}\right)(L_{13/14}^2 + \pi^2)\right].$$ (B.111)

Due to Eq. (B.105), the first two amplitudes in Eq. (B.106), as well as the first two in Eq. (B.111), can be obtained from each other by interchanging $1^+ \leftrightarrow 2^+$ which corresponds to $s_{13} \leftrightarrow s_{14}$ without an effect on the overall phase.

The amplitudes with only one or no gluon with negative helicity vanish at tree level,

$$A^{(0)}(1^+, 2^+, 3^+, 4^\pm) = A^{(0)}(1^+, 3^+, 4^\pm, 2^+) = A^{(0)}(1^+, 4^\pm, 2^+, 3^+) = 0,$$

$$B^{(0)}(1^+, 2^+, 3^+, 4^\pm) = B^{(0)}(1^+, 3^+, 4^\pm, 2^+) = B^{(0)}(1^+, 4^\pm, 2^+, 3^+) = 0.$$ (B.112)

The corresponding one-loop amplitudes are infrared finite. Those entering $\vec{C}_{+++-}$ are given by

$$A^{(1)}(1^+, 2^+, 3^+, 4^-) = 4\frac{[13]^2}{[41]\langle 12\rangle\langle 23\rangle[34]}\frac{1}{3}(C_A - 2T_F n_f)(s_{14} + s_{34})$$

453

$$= 4\, e^{\mathrm{i}\Phi_{+++-}} \frac{1}{3}(C_A - 2T_F\, n_f) \left(\frac{s_{13}}{s_{12}} + \frac{s_{13}}{s_{14}}\right),$$

$$A^{(1)}(1^+,3^+,4^-,2^+) = 4\frac{[23]^2}{[42]\langle 21\rangle\langle 13\rangle[34]}\frac{1}{3}(C_A - 2T_F\, n_f)\,(s_{13} + s_{12})$$

$$= 4\, e^{\mathrm{i}\Phi_{+++-}} \frac{1}{3}(C_A - 2T_F\, n_f) \left(\frac{s_{14}}{s_{12}} + \frac{s_{14}}{s_{13}}\right),$$

$$A^{(1)}(1^+,4^-,2^+,3^+) = 4\frac{[21]^2}{[42]\langle 23\rangle\langle 31\rangle[14]}\frac{1}{3}(C_A - 2T_F\, n_f)\,(s_{13} + s_{14})$$

$$= 4\, e^{\mathrm{i}\Phi_{+++-}} \frac{1}{3}(C_A - 2T_F\, n_f) \left(\frac{s_{12}}{s_{14}} + \frac{s_{12}}{s_{13}}\right),$$

$$B^{(1)}(1^+,2^+,3^+,4^-) = B^{(1)}(1^+,3^+,4^-,2^+) = B^{(1)}(1^+,4^-,2^+,3^+) = -16 T_F\, e^{\mathrm{i}\Phi_{+++-}}\,,$$

$$\text{(B.113)}$$

and those for $\vec{C}_{++++}$ are

$$A^{(1)}(1^+,2^+,3^+,4^+) = A^{(1)}(1^+,3^+,4^+,2^+) = A^{(1)}(1^+,4^+,2^+,3^+) = 4\, e^{\mathrm{i}\Phi_{++++}} \frac{1}{3}(C_A - 2T_F\, n_f)\,,$$

$$B^{(1)}(1^+,2^+,3^+,4^+) = B^{(1)}(1^+,3^+,4^+,2^+) = B^{(1)}(1^+,4^+,2^+,3^+) = 16 T_F\, e^{\mathrm{i}\Phi_{++++}}\,, \qquad \text{(B.114)}$$

where for convenience we have extracted the overall phases

$$e^{\mathrm{i}\Phi_{+++-}} = \frac{[12]}{\langle 12\rangle}\frac{[13]}{\langle 13\rangle}\frac{\langle 14\rangle}{[14]}\,, \qquad e^{\mathrm{i}\Phi_{++++}} = -\frac{[12]}{\langle 12\rangle}\frac{[34]}{\langle 34\rangle}\,. \qquad \text{(B.115)}$$

B.4.2 $pp \to 3$ Jets

In this appendix we give explicit expressions for all partial amplitudes that are required in Eqs. (5.209), (5.210), (5.217), and (5.226), for the various partonic channels for the $pp \to 3$ jets process. The one-loop amplitudes for these processes were calculated in Refs. [245, 246, 243], respectively. These chapters use $T_F = 1$ and $g_s T^a/\sqrt{2}$ for the $q\bar{q}g$ coupling. Thus, we can convert to our conventions by replacing $T^a \to \sqrt{2}T^a$, and identifying $1/N = C_A - 2C_F$ and $N = C_A$. Below we restrict ourselves to giving explicit expressions for the tree-level amplitudes, since the one-loop expressions are fairly lengthy.

For each partonic channel, we expand the amplitude as

$$X = [g_s(\mu)]^3 \sum_{n=0}^{\infty} X^{(n)} \left(\frac{\alpha_s(\mu)}{4\pi}\right)^n , \qquad (B.116)$$

where X stands for any of $A_{\text{div,fin}}$, $B_{\text{div,fin}}$.

$g q\bar{q}\, q'\bar{q}'$ and $g\, q\bar{q}\, q\bar{q}$

The tree-level amplitudes entering the Wilson coefficients in Eqs. (5.209) and (5.210) are given by

$$A^{(0)}(1^+; 2_q^+, 3_{\bar{q}}^-; 4_{q'}^+, 5_{\bar{q}'}^-) = \sqrt{2}\frac{\langle 25\rangle\langle 35\rangle^2}{\langle 12\rangle\langle 15\rangle\langle 23\rangle\langle 45\rangle} , \qquad A^{(0)}(1^+; 4_{q'}^+, 5_{\bar{q}'}^-; 2_q^+, 3_{\bar{q}}^-) = -\sqrt{2}\frac{\langle 35\rangle^2\langle 34\rangle}{\langle 13\rangle\langle 14\rangle\langle 23\rangle\langle 45\rangle} ,$$

$$A^{(0)}(1^+; 2_q^+, 3_{\bar{q}}^-; 4_{q'}^-, 5_{\bar{q}'}^+) = -\sqrt{2}\frac{\langle 25\rangle\langle 34\rangle^2}{\langle 12\rangle\langle 15\rangle\langle 23\rangle\langle 45\rangle} , \qquad A^{(0)}(1^+; 4_{q'}^-, 5_{\bar{q}'}^+; 2_q^+, 3_{\bar{q}}^-) = \sqrt{2}\frac{\langle 34\rangle^3}{\langle 13\rangle\langle 14\rangle\langle 23\rangle\langle 45\rangle} ,$$

$$B^{(0)}(1^+; 2_q^+, 3_{\bar{q}}^-; 4_{q'}^+, 5_{\bar{q}'}^-) = -\sqrt{2}\frac{\langle 23\rangle\langle 35\rangle^2}{\langle 12\rangle\langle 13\rangle\langle 23\rangle\langle 45\rangle} , \qquad B^{(0)}(1^+; 4_{q'}^+, 5_{\bar{q}'}^-; 2_q^+, 3_{\bar{q}}^-) = -\sqrt{2}\frac{\langle 35\rangle^2\langle 45\rangle}{\langle 14\rangle\langle 15\rangle\langle 23\rangle\langle 45\rangle} ,$$

$$B^{(0)}(1^+; 2_q^+, 3_{\bar{q}}^-; 4_{q'}^-, 5_{\bar{q}'}^+) - \sqrt{2}\frac{\langle 23\rangle\langle 34\rangle^2}{\langle 12\rangle\langle 13\rangle\langle 23\rangle\langle 45\rangle} , \qquad B^{(0)}(1^+; 4_{q'}^-, 5_{\bar{q}'}^+; 2_q^+, 3_{\bar{q}}^-) = \sqrt{2}\frac{\langle 34\rangle^2\langle 45\rangle}{\langle 14\rangle\langle 15\rangle\langle 23\rangle\langle 45\rangle} .$$

$$(B.117)$$

Of these helicity amplitudes only 4 are independent. The one-loop amplitudes were computed in Ref. [245].

$ggg\, q\bar{q}$

The three independent tree-level partial amplitudes which enter the Wilson coefficients in Eq. (5.217) are given by

$$A^{(0)}(1^+, 2^+, 3^-; 4_q^+, 5_{\bar{q}}^-) = 2\sqrt{2}\frac{\langle 34\rangle\langle 35\rangle^2}{\langle 12\rangle\langle 14\rangle\langle 23\rangle\langle 45\rangle} ,$$

$$A^{(0)}(2^+, 3^-, 1^+; 4_q^+, 5_{\bar{q}}^-) = -2\sqrt{2}\frac{\langle 34\rangle\langle 35\rangle^3}{\langle 13\rangle\langle 15\rangle\langle 23\rangle\langle 24\rangle\langle 45\rangle} ,$$

$$A^{(0)}(3^-, 1^+, 2^+; 4_q^+, 5_{\bar{q}}^-) = -2\sqrt{2}\frac{\langle 35\rangle^3}{\langle 12\rangle\langle 13\rangle\langle 25\rangle\langle 45\rangle} ,$$

$$B^{(0)} = C^{(0)} = 0 \,. \tag{B.118}$$

At tree level, the partial amplitudes for the other color structures vanish, $B^{(0)} = C^{(0)} = 0$. The one-loop amplitudes were computed in Ref. [246].

ggggg

The two independent partial amplitudes that enter the Wilson coefficients in Eq. (5.226) are given by the Parke-Taylor formula [493]

$$A^{(0)}(1^+, 2^+, 3^+, 4^-, 5^-) = 4\sqrt{2}\frac{\langle 45 \rangle^4}{\langle 12 \rangle \langle 23 \rangle \langle 34 \rangle \langle 45 \rangle \langle 51 \rangle}\,,$$

$$A^{(0)}(1^+, 2^+, 4^-, 3^+, 5^-) = 4\sqrt{2}\frac{\langle 45 \rangle^4}{\langle 12 \rangle \langle 15 \rangle \langle 24 \rangle \langle 34 \rangle \langle 35 \rangle}\,,$$

$$B^{(0)} = 0 \,. \tag{B.119}$$

All other amplitudes can be obtained by cyclic permutations. The double-trace color structure does not appear at tree level, so $B^{(0)} = 0$. The one-loop amplitudes were calculated in Ref. [243]. For completeness, I also add here a reference to my two neutrino papers [494, 495].

B.5   RGE Ingredients

In this appendix, we collect explicit results required for the running of the hard matching coefficients required to NNLL order. We expand the β function and cusp anomalous dimension in powers of α_s as

$$\beta(\alpha_s) = -2\alpha_s \sum_{n=0}^{\infty} \beta_n \left(\frac{\alpha_s}{4\pi}\right)^{n+1}, \qquad \Gamma_{\text{cusp}}(\alpha_s) = \sum_{n=0}^{\infty} \Gamma_n \left(\frac{\alpha_s}{4\pi}\right)^{n+1}. \tag{B.120}$$

Up to three-loop order in the $\overline{\text{MS}}$ scheme, the coefficients of the β function are [496, 497]

$$\beta_0 = \frac{11}{3} C_A - \frac{4}{3} T_F n_f\,, \qquad \beta_1 = \frac{34}{3} C_A^2 - \left(\frac{20}{3} C_A + 4 C_F\right) T_F n_f\,,$$

$$\beta_2 = \frac{2857}{54}\,C_A^3 + \left(C_F^2 - \frac{205}{18}\,C_F C_A - \frac{1415}{54}\,C_A^2\right)2T_F\,n_f + \left(\frac{11}{9}\,C_F + \frac{79}{54}\,C_A\right)4T_F^2\,n_f^2\,,$$

$$\text{(B.121)}$$

and for the cusp anomalous dimension they are [498, 499]

$$\Gamma_0 = 4\,, \qquad \Gamma_1 = \left(\frac{268}{9} - \frac{4\pi^2}{3}\right)C_A - \frac{80}{9}\,T_F\,n_f\,,$$

$$\Gamma_2 = \left(\frac{490}{3} - \frac{536\pi^2}{27} + \frac{44\pi^4}{45} + \frac{88\zeta_3}{3}\right)C_A^2 + \left(\frac{80\pi^2}{27} - \frac{836}{27} - \frac{112\zeta_3}{3}\right)C_A\,2T_F\,n_f$$

$$+ \left(32\zeta_3 - \frac{110}{3}\right)C_F\,2T_F\,n_f - \frac{64}{27}\,T_F^2\,n_f^2\,.$$

$$\text{(B.122)}$$

Note that here Γ_{cusp} does not include an overall color factor; it differs from the usual $q\bar{q}$ case by a factor of C_F.

For the noncusp anomalous dimension of the Wilson coefficient, which is color diagonal to two loops, we write

$$\widehat{\gamma}(\alpha_s) = [n_q\gamma_C^q(\alpha_s) + n_g\gamma_C^g(\alpha_s)]\mathbf{1} + \mathcal{O}(\alpha_s^3)\,, \tag{B.123}$$

as in Eq. (5.238). The quark and gluon noncusp anomalous dimensions,

$$\gamma_C^q(\alpha_s) = \left(\frac{\alpha_s}{4\pi}\right)\gamma_{C0}^q + \left(\frac{\alpha_s}{4\pi}\right)^2\gamma_{C1}^q\,, \qquad \gamma_C^g(\alpha_s) = \left(\frac{\alpha_s}{4\pi}\right)\gamma_{C0}^g + \left(\frac{\alpha_s}{4\pi}\right)^2\gamma_{C1}^g\,, \tag{B.124}$$

have the following coefficients

$$\gamma_{C0}^q = -3C_F\,,$$

$$\gamma_{C1}^q = -C_F\left[\left(\frac{41}{9} - 26\zeta_3\right)C_A + \left(\frac{3}{2} - 2\pi^2 + 24\zeta_3\right)C_F + \left(\frac{65}{18} + \frac{\pi^2}{2}\right)\beta_0\right]\,,$$

$$\gamma_{C0}^g = -\beta_0\,,$$

$$\gamma_{C1}^g = \left(-\frac{59}{9} + 2\zeta_3\right)C_A^2 + \left(-\frac{19}{9} + \frac{\pi^2}{6}\right)C_A\beta_0 - \beta_1\,. \tag{B.125}$$

The evolution kernels required for the resummation were defined in Eq. (5.237) by the

integrals

$$
K_\Gamma(\mu_0, \mu) = \int_{\alpha_s(\mu_0)}^{\alpha_s(\mu)} d\alpha_s \, \frac{\Gamma_{\text{cusp}}(\alpha_s)}{\beta(\alpha_s)} \int_{\alpha_s(\mu_0)}^{\alpha_s} d\alpha'_s \, \frac{1}{\beta(\alpha'_s)} \,,
$$

$$
\eta_\Gamma(\mu_0, \mu) = \int_{\alpha_s(\mu_0)}^{\alpha_s(\mu)} d\alpha_s \, \frac{\Gamma_{\text{cusp}}(\alpha_s)}{\beta(\alpha_s)} \,,
$$

$$
\widehat{K}_\gamma(\mu_0, \mu) = \int_{\alpha_s(\mu_0)}^{\alpha_s(\mu)} d\alpha_s \, \frac{\widehat{\gamma}(\alpha_s)}{\mathbf{t}(\alpha_s)} \,. \tag{B.126}
$$

Up to two loops, we can simplify the noncusp evolution kernel as

$$
\widehat{K}_\gamma(\mu_0, \mu) = \left(n_q K_\gamma^q(\mu_0, \mu) + n_g K_\gamma^g(\mu_0, \mu) \right) \mathbf{1} \,. \tag{B.127}
$$

Explicit results to NNLL order are given by

$$
\begin{aligned}
K_\Gamma(\mu_0, \mu) = {} & -\frac{\Gamma_0}{4\beta_0^2} \left\{ \frac{4\pi}{\alpha_s(\mu_0)} \left(1 - \frac{1}{r} - \ln r \right) + \left(\frac{\Gamma_1}{\Gamma_0} - \frac{\beta_1}{\beta_0} \right)(1 - r + \ln r) + \frac{\beta_1}{2\beta_0} \ln^2 r \right. \\
& \left. + \frac{\alpha_s(\mu_0)}{4\pi} \left[\left(\frac{\beta_1^2}{\beta_0^2} - \frac{\beta_2}{\beta_0} \right) \left(\frac{1 - r^2}{2} + \ln r \right) + \left(\frac{\beta_1 \Gamma_1}{\beta_0 \Gamma_0} - \frac{\beta_1^2}{\beta_0^2} \right)(1 - r + r \ln r) - \left(\frac{\Gamma_2}{\Gamma_0} - \frac{\beta_1 \Gamma_1}{\beta_0 \Gamma_0} \right) \frac{(1 - r)^2}{2} \right] \right\}, \\
\eta_\Gamma(\mu_0, \mu) = {} & -\frac{\Gamma_0}{2\beta_0} \left[\ln r + \frac{\alpha_s(\mu_0)}{4\pi} \left(\frac{\Gamma_1}{\Gamma_0} - \frac{\beta_1}{\beta_0} \right)(r - 1) + \frac{\alpha_s^2(\mu_0)}{16\pi^2} \left(\frac{\Gamma_2}{\Gamma_0} - \frac{\beta_1 \Gamma_1}{\beta_0 \Gamma_0} + \frac{\beta_1^2}{\beta_0^2} - \frac{\beta_2}{\beta_0} \right) \frac{r^2 - 1}{2} \right], \\
K_\gamma^q(\mu_0, \mu) = {} & -\frac{\gamma_{C0}^q}{2\beta_0} \left[\ln r + \frac{\alpha_s(\mu_0)}{4\pi} \left(\frac{\gamma_{C1}^q}{\gamma_{C0}^q} - \frac{\beta_1}{\beta_0} \right)(r - 1) \right], \\
K_\gamma^g(\mu_0, \mu) = {} & -\frac{\gamma_{C0}^g}{2\beta_0} \left[\ln r + \frac{\alpha_s(\mu_0)}{4\pi} \left(\frac{\gamma_{C1}^g}{\gamma_{C0}^g} - \frac{\beta_1}{\beta_0} \right)(r - 1) \right], \tag{B.128}
\end{aligned}
$$

with $r = \alpha_s(\mu)/\alpha_s(\mu_0)$. The running coupling in the above equations is given by the three-loop expression

$$
\frac{1}{\alpha_s(\mu)} = \frac{X}{\alpha_s(\mu_0)} + \frac{\beta_1}{4\pi\beta_0} \ln X + \frac{\alpha_s(\mu_0)}{16\pi^2} \left[\frac{\beta_2}{\beta_0} \left(1 - \frac{1}{X} \right) + \frac{\beta_1^2}{\beta_0^2} \left(\frac{\ln X}{X} + \frac{1}{X} - 1 \right) \right], \tag{B.129}
$$

with $X \equiv 1 + \alpha_s(\mu_0) \beta_0 \ln(\mu/\mu_0)/(2\pi)$.

B.6 Color Sum Matrices

For each specific process considered in the text we decomposed the Wilson coefficients in a color basis as

$$C^{a_1\cdots a_n}_{+\cdots(\cdots-)} = \sum_k C^k_{+\cdots(\cdots-)} T^{a_1\cdots a_n}_k \equiv \bar{T}^{a_1\cdots a_n} \vec{C}_{+\cdots(\cdots-)}\,, \tag{B.130}$$

where $\bar{T}^{a_1\cdots a_n}$ is a row vector of color structures which form a complete basis of the allowed color structures for the particular process. Since convenient color bases are generically not orthogonal, the scalar product between Wilson coefficients is nontrivial. The $\vec{C}^\dagger$ is given by

$$\vec{C}^\dagger = [C^{a_1\cdots a_n}]^* \, \bar{T}^{a_1\cdots a_n} = \vec{C}^{*T} \widehat{T}\,, \tag{B.131}$$

where

$$\widehat{T} = \sum_{a_1,\ldots,a_n} (\bar{T}^{a_1\cdots a_n})^\dagger \bar{T}^{a_1\cdots a_n} \tag{B.132}$$

is the matrix of color sums.

In this appendix we give explicit expressions for $\widehat{T}$ for all the processes in this chapter, both for general $SU(N)$, as well as a numerical result for the specific case of $N = 3$. For simplicity, in this section we restrict ourselves to the normalization convention $T_F = 1/2$, and $C_A = N$, and write the results for general $SU(N)$ in terms of only C_A and C_F.

For $q\bar{q}$ and gg in the basis in Eq. (5.105), we have

$$\widehat{T}_{q\bar{q}} = C_A = 3\,, \qquad \widehat{T}_{gg} = 2C_A C_F = 8\,. \tag{B.133}$$

For $g\,q\bar{q}$ and ggg in the basis Eq. (5.110), we have

$$\widehat{T}_{g\,q\bar{q}} = C_A C_F = 4\,,$$

$$\widehat{T}_{ggg} = 2C_F \begin{pmatrix} C_A^2 & 0 \\ 0 & C_A^2 - 4 \end{pmatrix} = \frac{8}{3}\begin{pmatrix} 9 & 0 \\ 0 & 5 \end{pmatrix}\,. \tag{B.134}$$

For $q\bar{q}\,q\bar{q}$ and $q\bar{q}\,q'\bar{q}'$ in the basis Eq. (5.186), we have

$$\widehat{T}_{q\bar{q}q\bar{q}} = \widehat{T}_{q\bar{q}q'\bar{q}'} = \begin{pmatrix} C_A^2 & C_A \\ C_A & C_A^2 \end{pmatrix} = \begin{pmatrix} 9 & 3 \\ 3 & 9 \end{pmatrix}. \tag{B.135}$$

For $gg\,q\bar{q}$ in the basis Eq. (5.192), we have

$$\widehat{T}_{ggq\bar{q}} = \frac{C_A C_F}{2} \begin{pmatrix} 2C_F & 2C_F - C_A & 1 \\ 2C_F - C_A & 2C_F & 1 \\ 1 & 1 & C_A \end{pmatrix}$$

$$= \frac{2}{3} \begin{pmatrix} 8 & -1 & 3 \\ -1 & 8 & 3 \\ 3 & 3 & 9 \end{pmatrix}, \tag{B.136}$$

and for $gggg$ in the basis Eq. (5.197), we have

$$\widehat{T}_{gggg} = \frac{C_A C_F}{4} \begin{pmatrix} a & b & b & c & d & c \\ b & a & b & c & c & d \\ b & b & a & d & c & c \\ c & c & d & e & f & f \\ d & c & c & f & e & f \\ c & d & c & f & f & e \end{pmatrix}, \tag{B.137}$$

where

$$a = C_A^2 - \frac{9}{2}C_A C_F + 6C_F^2 + \frac{1}{4} = \frac{23}{12},$$
$$b = C_A^2 - 5C_A C_F + 6C_F^2 = -\frac{1}{3},$$
$$c = C_F = \frac{4}{3}, \qquad d = \frac{(2C_F - C_A)}{2} = -\frac{1}{6},$$
$$e = C_F C_A = 4, \qquad f = \frac{1}{2}. \tag{B.138}$$

For $g\,q\bar{q}\,q\bar{q}$ and $g\,q\bar{q}\,q'\bar{q}'$ in the basis Eq. (5.206) we have

$$\widehat{T}_{g\,q\bar{q}\,q\bar{q}} = C_A C_F \begin{pmatrix} C_A & 0 & 1 & 1 \\ 0 & C_A & 1 & 1 \\ 1 & 1 & C_A & 0 \\ 1 & 1 & 0 & C_A \end{pmatrix}$$

$$= 4 \begin{pmatrix} 3 & 0 & 1 & 1 \\ 0 & 3 & 1 & 1 \\ 1 & 1 & 3 & 0 \\ 1 & 1 & 0 & 3 \end{pmatrix} . \tag{B.139}$$

For $ggg\,q\bar{q}$ in the basis Eq. (5.215) we have

$$\widehat{T}_{ggg\,q\bar{q}} = \frac{C_F}{4} \begin{pmatrix} a & b & b & c & d & d & e & f & f & i & j \\ b & a & b & d & c & d & f & e & f & i & j \\ b & b & a & d & d & c & f & f & e & i & j \\ c & d & d & a & b & b & e & f & f & j & i \\ d & c & d & b & a & b & f & e & f & j & i \\ d & d & c & b & b & a & f & f & e & j & i \\ e & f & f & e & f & f & g & h & h & 0 & 0 \\ f & e & f & f & e & f & h & g & h & 0 & 0 \\ f & f & e & f & f & e & h & h & g & 0 & 0 \\ i & i & i & j & j & j & 0 & 0 & 0 & i & j \\ j & j & j & i & i & i & 0 & 0 & 0 & j & i \end{pmatrix} , \tag{B.140}$$

where

$$a = 4C_A C_F^2 = \frac{64}{3} , \qquad b = C_A - 2C_F = \frac{1}{3} ,$$
$$c = (C_A^2 + 1)(C_A - 2C_F) = \frac{10}{3} , \qquad d = -2C_F = -\frac{8}{3} ,$$

461

$$e = -1, \qquad f = 2C_A C_F = 8, \qquad g = 2C_A^2 C_F = 24,$$

$$h = C_A = 3, \qquad i = C_A^2 - 2 = 7, \qquad j = -2.$$ (B.141)

For $ggggg$ in the basis Eq. (5.221) we have

$$\widehat{T}_{ggggg} = \frac{C_F}{32} \begin{pmatrix} \widehat{X}_1 & \widehat{X}_2 \\ \widehat{X}_2^T & \widehat{X}_3 \end{pmatrix},$$ (B.142)

where

$$\widehat{X}_1 = \begin{pmatrix}
a & -b-c-b & c & -b-c & b & -c-b-c & 0 \\
-b & a & b-c & b & c & -b-c & b & -c & 0 & c \\
-c & b & a & -b-c-b-c-b-c & 0 & c & -b \\
-b-c-b & a & b & -c & b & -c & 0 & c & b & c \\
c & b & -c & b & a & -b-c & 0 & c & -b & c & b \\
-b & c & -b-c-b & a & 0 & c & b & c & -b & c \\
-c-b-c & b & -c & 0 & a & -b & c & -b-c-b \\
b & -c-b-c & 0 & c & -b & a & b & c & b & -c \\
-c & b & -c & 0 & c & b & c & b & a & -b & c & -b \\
-b-c & 0 & c & -b & c & -b & c & -b & a & b & c \\
-c & 0 & c & b & c & -b-c & b & c & b & a & -b \\
0 & c & -b & c & b & c & -b-c-b & c & -b & a
\end{pmatrix},$$

462

$$\widehat{X}_2 = \begin{pmatrix}
-d & d & -e & e & e & -e & -d & -d & -d & e \\
-d & -d & d & e & -e & -e & e & e & d & -d \\
-e & -e & d & d & e & -e & -e & d & d & d \\
d & -e & -e & -d & -d & -e & e & d & -d & -e \\
-d & -e & e & -e & -d & -d & -e & e & -d & -d \\
e & -e & -e & -e & e & d & d & -d & d & -d \\
d & -d & -d & -d & d & -e & -e & e & -e & e \\
-e & d & -d & d & -e & -e & d & -d & -e & -e \\
-e & -d & -e & e & e & -d & d & -e & e & -d \\
d & d & e & e & -d & d & d & e & e & e \\
e & e & -e & d & -d & -d & d & d & -e & e \\
-e & e & d & -d & -d & d & -e & -e & -e & -d
\end{pmatrix} ,$$

$$\widehat{X}_3 = \begin{pmatrix}
f & 0 & -g & -g & 0 & g & g & g & 0 & g \\
0 & f & 0 & g & -g & g & 0 & -g & -g & g \\
-g & 0 & f & 0 & -g & g & -g & g & g & 0 \\
-g & g & 0 & f & 0 & -g & g & 0 & g & g \\
0 & -g & -g & 0 & f & 0 & -g & -g & g & g \\
g & g & g & -g & 0 & f & 0 & -g & g & 0 \\
g & 0 & -g & g & -g & 0 & f & 0 & g & -g \\
g & -g & g & 0 & -g & -g & 0 & f & 0 & g \\
0 & -g & g & g & g & g & g & 0 & f & 0 \\
g & g & 0 & g & g & 0 & -g & g & 0 & f
\end{pmatrix} , \qquad (\text{B.143})$$

and

$$a = C_A^4 - 4C_A^2 + 10 = 55 , \qquad b = 2C_A^2 - 4 = 14 ,$$
$$c = 2 , \qquad d = 2C_A^2 C_F = 24 , \qquad e = C_A = 3 ,$$

$$f = 2C_A^3 C_F = 72\,, \qquad g = C_A^2 = 9\,. \tag{B.144}$$

B.7 IR Divergences

In this appendix, we explicitly check that the IR divergences of QCD are reproduced by SCET. This ensures that they drop out in the one-loop matching, and that the resulting Wilson coefficients are IR finite. They also provide a very useful cross check when converting from the different conventions used in the literature to ours.

The one-loop matching equation relating the SCET operators and their Wilson coefficients to the QCD amplitude is

$$\langle \vec{O}^\dagger \rangle^{(0)} \vec{C}^{(1)} + \langle \vec{O}^\dagger \rangle^{(1)} \vec{C}^{(0)} = -i\mathcal{A}^{(1)}\,. \tag{B.145}$$

First we determine the residues of the propagators entering the LSZ reduction formula. Regulating both UV and IR divergences in dimensional regularization, all bare loop integrals in SCET are scaleless and vanish, i.e. the UV and IR divergences cancel. In particular, for the self-energy diagrams, we have

$$\Sigma = \Sigma_{\mathrm{UV}} + \Sigma_{\mathrm{IR}} = 0\,. \tag{B.146}$$

The UV divergences Σ_{UV} plus possible additional UV finite terms Σ_x (as dictated by the renormalization scheme) determine the wave function renormalization Z_ξ. The remainder $\Sigma_{\mathrm{IR}} - \Sigma_x$ enters the residue R_ξ

$$\begin{aligned} Z_\xi^{-1} &= 1 - \frac{\mathrm{d}(\Sigma_{\mathrm{UV}} + \Sigma_x)}{\mathrm{d}\not{p}}\bigg|_{\not{p}=0}\,, \\ R_\xi^{-1} &= 1 - \frac{\mathrm{d}(\Sigma_{\mathrm{IR}} - \Sigma_x)}{\mathrm{d}\not{p}}\bigg|_{\not{p}=0}\,. \end{aligned} \tag{B.147}$$

At one loop in pure dimensional regularization, we then have $R_\xi = Z_\xi^{-1}$, and similarly for gluons $R_A = Z_A^{-1}$. In the on-shell scheme $\Sigma_x = \Sigma_{\mathrm{IR}}$, so with pure dimensional regularization

$$Z_\xi = R_\xi = Z_A = R_A = 1.$$

Since all loop diagrams contributing to $\langle \vec{O}^\dagger \rangle^{(1)}$ vanish, the only nonzero contributions come from the counterterm in Eq. (5.229) and the one-loop residues. At one loop we find

$$
\begin{aligned}
\langle \vec{O}^\dagger \rangle^{(1)} \vec{C}^{(0)} &= \langle \vec{O}^\dagger \rangle^{(0)} \Big[\big(Z_\xi^{n_q/2} Z_A^{n_g/2} \widehat{Z}_C - 1 \big) + \big(R_\xi^{n_q/2} R_A^{n_g/2} - 1 \big) \Big] \vec{C}^{(0)} = \langle \vec{O}^\dagger \rangle^{(0)} (\widehat{Z}_C - 1) \vec{C}^{(0)} \\
&= \langle \vec{O}^\dagger \rangle^{(0)} \frac{\alpha_s}{4\pi} \Big[-\frac{1}{\epsilon^2} \big(n_g C_A + n_q C_F \big) + \frac{1}{\epsilon} \Big(-\frac{1}{2} n_g \beta_0 - \frac{3}{2} n_q C_F + 2\widehat{\Delta}(\mu^2) \Big) \Big] \vec{C}^{(0)} \,,
\end{aligned}
$$

$$(B.148)$$

where we used the explicit expression for $\widehat{Z}_C$ derived in Sec. 5.8.2. One can easily check that this exactly reproduces the IR-divergent parts of the QCD amplitudes. For example, for $gg\, q\bar{q}$, we have

$$
\Big[-\frac{1}{\epsilon^2}(2C_A + 2C_F) + \frac{1}{\epsilon}\Big(-\beta_0 - 3C_F + 2\widehat{\Delta}_{gg\, q\bar{q}}(\mu^2) \Big) \Big] C^{'(0)}_{+-(+)}(p_1, p_2; p_3, p_4) =
\begin{pmatrix} A^{(1)}_{\mathrm{div}}(1^+, 2^-; 3_q^+, 4_{\bar{q}}^-) \\ A^{(1)}_{\mathrm{div}}(2^-, 1^+; 3_q^+, 4_{\bar{q}}^-) \\ B^{(1)}_{\mathrm{div}}(1^+, 2^-; 3_q^+, 4_{\bar{q}}^-) \end{pmatrix} .
$$

$$(B.149)$$

Hence, the IR divergences in $\langle \vec{O}^\dagger \rangle^{(1)} \vec{C}^{(0)}$ and $\mathcal{A}^{(1)}$ cancel each other and do not enter in $\vec{C}^{(1)}$, as must be the case.

Bibliography

[1] A. H. Hoang, D. W. Kolodrubetz, V. Mateu, and I. W. Stewart, *C-parameter Distribution at N^3LL' including Power Corrections*, arXiv:1411.6633.

[2] R. Abbate, M. Fickinger, A. H. Hoang, V. Mateu, and I. W. Stewart, *Thrust at N^3LL with Power Corrections and a Precision Global Fit for alphas(mZ)*, Phys. Rev. D **83** (2011) 074021, [arXiv:1006.3080].

[3] T. A. collaboration, *Search for resonances with boson-tagged jets in 3.2 fbâĽŠ1 of p p collisions at âĽŽ s = 13 TeV collected with the ATLAS detector*, .

[4] T. A. collaboration, *Studies of b-tagging performance and jet substructure in a high p_T $g \to b\bar{b}$ rich sample of large-R jets from pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, .

[5] A. J. Larkoski, I. Moult, and D. Neill, *Power Counting to Better Jet Observables*, JHEP **1412** (2014) 009, [arXiv:1409.6298].

[6] **ATLAS** Collaboration, G. Aad et al., *Identification of Boosted, Hadronically Decaying W Bosons and Comparisons with ATLAS Data Taken at $\sqrt{s} = 8$ TeV*, arXiv:1510.05821.

[7] G. Aad et al., *Performance of jet substructure techniques in early $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, ATLAS-CONF-2015-035 (2015), no. ATLAS-CONF-2015-035.

[8] A. J. Larkoski, I. Moult, and D. Neill, *Building a Better Boosted Top Tagger*, arXiv:1411.0665.

[9] A. J. Larkoski and I. Moult, *The Singular Behavior of Jet Substructure Observables*, Phys. Rev. **D93** (2016) 014017, [arXiv:1510.08459].

[10] A. J. Larkoski, I. Moult, and D. Neill, *Toward Multi-Differential Cross Sections: Measuring Two Angularities on a Single Jet*, arXiv:1401.4458.

[11] A. J. Larkoski, I. Moult, and D. Neill, *Non-Global Logarithms, Factorization, and the Soft Substructure of Jets*, arXiv:1501.04596.

467

[12] A. J. Larkoski, I. Moult, and D. Neill, *Analytic Boosted Boson Discrimination*, arXiv:1507.03018.

[13] I. W. Stewart, F. J. Tackmann, J. R. Walsh, and S. Zuberi, *Jet p_T resummation in Higgs production at $NNLL' + NNLO$*, Phys. Rev. **D89** (2014), no. 5 054001, [arXiv:1307.1808].

[14] I. Moult and I. W. Stewart, *Jet Vetoes interfering with $H \to WW$*, JHEP **1409** (2014) 129, [arXiv:1405.5534].

[15] **CMS** Collaboration, V. Khachatryan et al., *Constraints on the Higgs boson width from off-shell production and decay to Z-boson pairs*, Phys. Lett. **B736** (2014) 64–85, [arXiv:1405.3455].

[16] J. C. Collins, D. E. Soper, and G. Sterman, *Soft gluons and factorization*, Nucl. Phys. *B* **308** (1988) 833.

[17] J. R. Gaunt, *Glauber Gluons and Multiple Parton Interactions*, JHEP **07** (2014) 110, [arXiv:1405.2080].

[18] M. Zeng, *Drell-Yan process with jet vetoes: breaking of generalized factorization*, arXiv:1507.01652.

[19] I. Z. Rothstein and I. W. Stewart, *An Effective Field Theory for Forward Scattering and Factorization Violation*, arXiv:1601.04695.

[20] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *Factorization at the LHC: From PDFs to Initial State Jets*, Phys. Rev. *D* **81** (2010) 094035, [arXiv:0910.0467].

[21] C. W. Bauer, S. Fleming, and M. E. Luke, *Summing Sudakov logarithms in B -> X(s gamma) in effective field theory*, Phys.Rev. **D63** (2000) 014006, [hep-ph/0005275].

[22] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, *An effective field theory for collinear and soft gluons: Heavy to light decays*, Phys. Rev. *D* **63** (2001) 114020, [hep-ph/0011336].

[23] C. W. Bauer and I. W. Stewart, *Invariant operators in collinear effective theory*, Phys.Lett. **B516** (2001) 134–142, [hep-ph/0107001].

[24] C. W. Bauer, D. Pirjol, and I. W. Stewart, *Soft-collinear factorization in effective field theory*, Phys. Rev. *D* **65** (2002) 054022, [hep-ph/0109045].

[25] C. W. Bauer, S. Fleming, D. Pirjol, I. Z. Rothstein, and I. W. Stewart, *Hard scattering factorization from effective field theory*, Phys. Rev. *D* **66** (2002) 014017, [hep-ph/0202088].

[26] A. V. Manohar, T. Mehen, D. Pirjol, and I. W. Stewart, *Reparameterization invariance for collinear operators*, Phys. Lett. *B* **539** (2002) 59–66, [hep-ph/0204229].

[27] J. Chay and C. Kim, *Collinear effective theory at subleading order and its application to heavy-light currents*, Phys. Rev. D **65** (2002) 114016, [hep-ph/0201197].

[28] C. W. Bauer, D. Pirjol, and I. W. Stewart, *Factorization and endpoint singularities in heavy to light decays*, Phys.Rev. **D67** (2003) 071502, [hep-ph/0211069].

[29] C. Marcantonini and I. W. Stewart, *Reparameterization Invariant Collinear Operators*, Phys. Rev. **D79** (2009) 065028, [arXiv:0809.1093].

[30] A. J. Larkoski, G. P. Salam, and J. Thaler, *Energy Correlation Functions for Jet Substructure*, JHEP **1306** (2013) 108, [arXiv:1305.0007].

[31] J. Thaler and K. Van Tilburg, *Maximizing Boosted Top Identification by Minimizing N-subjettiness*, JHEP **1202** (2012) 093, [arXiv:1108.2701].

[32] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, JHEP **1103** (2011) 015, [arXiv:1011.2268].

[33] A. J. Larkoski, D. Neill, and J. Thaler, *Jet Shapes with the Broadening Axis*, JHEP **1404** (2014) 017, [arXiv:1401.2158].

[34] M. Procura, W. J. Waalewijn, and L. Zeune, *Resummation of Double-Differential Cross Sections and Fully-Unintegrated Parton Distribution Functions*, JHEP **1502** (2015) 117, [arXiv:1410.6483].

[35] Y.-T. Chien, A. Hornig, and C. Lee, *Soft-collinear mode for jet cross sections in soft collinear effective theory*, Phys. Rev. **D93** (2016), no. 1 014033, [arXiv:1509.04287].

[36] P. Pietrulewicz, F. J. Tackmann, and W. J. Waalewijn, *Factorization and Resummation for Generic Hierarchies between Jets*, arXiv:1601.05088.

[37] A. J. Larkoski, J. Thaler, and W. J. Waalewijn, *Gaining (Mutual) Information about Quark/Gluon Discrimination*, arXiv:1408.3122.

[38] M. Dasgupta and G. Salam, *Resummation of nonglobal QCD observables*, Phys.Lett. **B512** (2001) 323–330, [hep-ph/0104277].

[39] A. J. Larkoski and I. Moult, *Nonglobal correlations in collider physics*, Phys. Rev. **D93** (2016), no. 1 014012, [arXiv:1510.05657].

[40] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, *Soft Drop*, JHEP **1405** (2014) 146, [arXiv:1402.2657].

[41] C. Frye, A. J. Larkoski, M. D. Schwartz, and K. Yan, *Precision physics with pile-up insensitive observables*, arXiv:1603.06375.

[42] C. Frye, A. J. Larkoski, M. D. Schwartz, and K. Yan, *Factorization for groomed jet substructure beyond the next-to-leading logarithm*, arXiv:1603.09338.

[43] I. Moult, I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *Employing Helicity Amplitudes for Resummation*, arXiv:1508.02397.

[44] D. W. Kolodrubetz, I. Moult, and I. W. Stewart, *Building Blocks for Subleading Helicity Operators*, arXiv:1601.02607.

[45] F. Caola and K. Melnikov, *Constraining the Higgs boson width with ZZ production at the LHC*, Phys.Rev. **D88** (2013) 054024, [arXiv:1307.4935].

[46] A. Abdesselam, E. B. Kuutmann, U. Bitenc, G. Brooijmans, J. Butterworth, et al., *Boosted objects: A Probe of beyond the Standard Model physics*, Eur.Phys.J. **C71** (2011) 1661, [arXiv:1012.5412].

[47] A. Altheimer, S. Arora, L. Asquith, G. Brooijmans, J. Butterworth, et al., *Jet Substructure at the Tevatron and LHC: New results, new tools, new benchmarks*, J.Phys. **G39** (2012) 063001, [arXiv:1201.0008].

[48] A. Altheimer, A. Arce, L. Asquith, J. Backus Mayes, E. Bergeaas Kuutmann, et al., *Boosted objects and jet substructure at the LHC. Report of BOOST2012, held at IFIC Valencia, 23rd-27th of July 2012*, Eur.Phys.J. **C74** (2014) 2792, [arXiv:1311.2708].

[49] **CMS Collaboration** Collaboration, *Jet Substructure Algorithms*, .

[50] **ATLAS Collaboration** Collaboration, D. W. Miller, *Jet substructure in ATLAS*, Tech. Rep. ATL-PHYS-PROC-2011-142, 2011.

[51] **ATLAS Collaboration** Collaboration, *Studies of the impact and mitigation of pile-up on large-r and groomed jets in atlas at $\sqrt{s} = 7$ tev*, Tech. Rep. ATLAS-CONF-2012-066, CERN, Geneva, Jul, 2012.

[52] **CMS Collaboration** Collaboration, S. Chatrchyan et al., *Shape, transverse size, and charged hadron multiplicity of jets in pp collisions at 7 TeV*, JHEP **1206** (2012) 160, [arXiv:1204.3170].

[53] **ATLAS Collaboration** Collaboration, *Studies of the impact and mitigation of pile-up on large-R and groomed jets in ATLAS at $\sqrt{s} = 7$ TeV*, .

[54] **ATLAS Collaboration** Collaboration, G. Aad et al., *ATLAS measurements of the properties of jets for boosted particle searches*, Phys.Rev. **D86** (2012) 072006, [arXiv:1206.5369].

[55] **ATLAS Collaboration** Collaboration, *Performance of large-R jets and jet substructure reconstruction with the ATLAS detector*, .

[56] **ATLAS Collaboration** Collaboration, G. Aad et al., *Jet mass and substructure of inclusive jets in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS experiment*, JHEP **1205** (2012) 128, [arXiv:1203.4606].

470

[57] **ATLAS Collaboration** Collaboration, G. Aad et al., *Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector*, JHEP **1309** (2013) 076, [arXiv:1306.4945].

[58] **ATLAS Collaboration** Collaboration, G. Aad et al., *Measurement of jet shapes in top pair events at sqrt(s)* = 7 TeV using the ATLAS detector, arXiv:1307.5749.

[59] **ATLAS Collaboration** Collaboration, *Performance and Validation of Q-Jets at the ATLAS Detector in pp Collisions at $\sqrt{s}$=8 TeV in 2012*, Tech. Rep. ATLAS-CONF-2013-087, ATLAS-COM-CONF-2013-099, 2013.

[60] **ATLAS Collaboration** Collaboration, *Jet Charge Studies with the ATLAS Detector Using $\sqrt{s}$ = 8 TeV Proton-Proton Collision Data*, Tech. Rep. ATLAS-CONF-2013-086, ATLAS-COM-CONF-2013-101, 2013.

[61] **ATLAS Collaboration** Collaboration, *Performance of pile-up subtraction for jet shapes*, Tech. Rep. ATLAS-CONF-2013-085, ATLAS-COM-CONF-2013-100, 2013.

[62] **ATLAS Collaboration** Collaboration, *Pile-up subtraction and suppression for jets in ATLAS*, Tech. Rep. ATLAS-CONF-2013-083, ATLAS-COM-CONF-2013-097, 2013.

[63] **CMS Collaboration** Collaboration, C. Collaboration, *Identifying Hadronically Decaying Vector Bosons Merged into a Single Jet*, .

[64] **CMS Collaboration** Collaboration, C. Collaboration, *Performance of quark/gluon discrimination in 8 TeV pp data*, Tech. Rep. CMS-PAS-JME-13-002, 2013.

[65] **CMS Collaboration** Collaboration, C. Collaboration, *Pileup Jet Identification*, Tech. Rep. CMS-PAS-JME-13-005, 2013.

[66] **CMS Collaboration** Collaboration, *Jet substructure algorithms*, Tech. Rep. CMS-PAS-JME-10-013, CERN, Geneva, 2011.

[67] **CMS Collaboration** Collaboration, *Measurement of the subjet multiplicity in dijet events from proton-proton collisions at sqrt(s)* = 7 tev, Tech. Rep. CMS-PAS-QCD-10-041, CERN, Geneva, 2010.

[68] **ATLAS Collaboration** Collaboration, G. Aad et al., *Light-quark and gluon jet discrimination in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, arXiv:1405.6583.

[69] **ATLAS** Collaboration, P. Loch, *Studies of jet shapes and jet substructure in proton-proton collisions at $\sqrt{s}$ = 7 TeV with ATLAS*, PoS **EPS-HEP2013** (2013) 442.

[70] **CMS Collaboration** Collaboration, C. Collaboration, *Boosted Top Jet Tagging at CMS*, .

[71] **CMS Collaboration** Collaboration, *Search for BSM ttbar Production in the Boosted All-Hadronic Final State*, Tech. Rep. CMS-PAS-EXO-11-006, 2011.

[72] **ATLAS, CMS Collaboration** Collaboration, S. Fleischmann, *Boosted top quark techniques and searches for $t\bar{t}$ resonances at the LHC*, J.Phys.Conf.Ser. **452** (2013), no. 1 012034.

[73] **ATLAS, CMS Collaboration** Collaboration, J. Pilot, *Boosted Top Quarks, Top Pair Resonances, and Top Partner Searches at the LHC*, EPJ Web Conf. **60** (2013) 09003.

[74] **ATLAS Collaboration** Collaboration, *Performance of boosted top quark identification in 2012 ATLAS data*, Tech. Rep. ATLAS-CONF-2013-084, ATLAS-COM-CONF-2013-074, 2013.

[75] **CMS Collaboration** Collaboration, S. Chatrchyan et al., *Search for Anomalous $t\bar{t}$ Production in the Highly-Boosted All-Hadronic Final State*, JHEP **1209** (2012) 029, [arXiv:1204.2488].

[76] **CMS Collaboration** Collaboration, S. Chatrchyan et al., *Search for a Higgs boson in the decay channel H to ZZ(*) to q qbar ℓ^- l+ in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP **1204** (2012) 036, [arXiv:1202.1416].

[77] **CMS Collaboration** Collaboration, C. Collaboration, *Search for a Standard Model-like Higgs boson decaying into WW to l nu qqbar in pp collisions at sqrt s = 8 TeV*, .

[78] **CMS Collaboration** Collaboration, C. Collaboration, *Search for pair-produced vector-like quarks of charge -1/3 decaying to bH using boosted Higgs jet-tagging in pp collisions at sqrt(s) = 8 TeV*, .

[79] **CMS Collaboration** Collaboration, C. Collaboration, *Search for top-Higgs resonances in all-hadronic final states using jet substructure methods*, .

[80] S. Catani, G. Turnock, and B. Webber, *Heavy jet mass distribution in e+ e- annihilation*, Phys.Lett. **B272** (1991) 368–372.

[81] Y.-T. Chien and M. D. Schwartz, *Resummation of heavy jet mass and comparison to LEP data*, JHEP **1008** (2010) 058, [arXiv:1005.1644].

[82] Y.-T. Chien, R. Kelley, M. D. Schwartz, and H. X. Zhu, *Resummation of Jet Mass at Hadron Colliders*, Phys.Rev. **D87** (2013) 014010, [arXiv:1208.0010].

[83] M. Dasgupta, K. Khelifa-Kerfa, S. Marzani, and M. Spannowsky, *On jet mass distributions in Z+jet and dijet processes at the LHC*, JHEP **1210** (2012) 126, [arXiv:1207.1640].

[84] T. T. Jouttenus, I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *Jet Mass Spectra in Higgs + One Jet at NNLL*, Phys. Rev. D **88** (2013) 054031, [arXiv:1302.0846].

[85] I. Feige, M. D. Schwartz, I. W. Stewart, and J. Thaler, *Precision Jet Substructure from Boosted Event Shapes*, Phys.Rev.Lett. **109** (2012) 092001, [arXiv:1204.3898].

[86] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, *Towards an understanding of jet substructure*, JHEP **1309** (2013) 029, [arXiv:1307.0007].

[87] M. Dasgupta, A. Fregoso, S. Marzani, and A. Powling, *Jet substructure with analytical methods*, Eur.Phys.J. **C73** (2013), no. 11 2623, [arXiv:1307.0013].

[88] A. J. Larkoski and J. Thaler, *Unsafe but Calculable: Ratios of Angularities in Perturbative QCD*, JHEP **1309** (2013) 137, [arXiv:1307.1699].

[89] A. J. Larkoski and J. Thaler, *Aspects of Jets at 100 TeV*, arXiv:1406.7011.

[90] J. R. Walsh and S. Zuberi, *Factorization Constraints on Jet Substructure*, arXiv:1110.5333.

[91] S. Catani, G. Turnock, and B. Webber, *Jet broadening measures in e^+e^- annihilation*, Phys.Lett. **B295** (1992) 269–276.

[92] Y. L. Dokshitzer, A. Lucenti, G. Marchesini, and G. Salam, *On the QCD analysis of jet broadening*, JHEP **9801** (1998) 011, [hep-ph/9801324].

[93] A. Banfi, G. P. Salam, and G. Zanderighi, *Principles of general final-state resummation and automated implementation*, JHEP **0503** (2005) 073, [hep-ph/0407286].

[94] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, [arXiv:0802.1189].

[95] D. Bertolini, T. Chan, and J. Thaler, *Jet Observables Without Jet Algorithms*, JHEP **1404** (2014) 013, [arXiv:1310.7584].

[96] G. Salam, *Unpublished, Unpublished.*

[97] G. C. Blazey, J. R. Dittmann, S. D. Ellis, V. D. Elvira, K. Frame, et al., *Run II jet physics*, hep-ex/0005012.

[98] P. E. Rakow and B. Webber, *Transverse Momentum Moments of Hadron Distributions in QCD Jets*, Nucl.Phys. **B191** (1981) 63.

473

[99] R. K. Ellis and B. Webber, *QCD Jet Broadening in Hadron Hadron Collisions,* *Conf.Proc.* **C860623** (1986) 74.

[100] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual,* *Eur. Phys. J. C* **72** (2012) 1896, [arXiv:1111.6097].

[101] "Fastjet contrib." http://fastjet.hepforge.org/contrib/.

[102] M. Beneke and V. A. Smirnov, *Asymptotic expansion of Feynman integrals near threshold,* *Nucl.Phys.* **B522** (1998) 321–344, [hep-ph/9711391].

[103] V. A. Smirnov, *'Strategy of regions': Expansions of Feynman diagrams both in Euclidean and pseudo-Euclidean regimes,* hep-ph/0101152.

[104] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *N-Jettiness: An Inclusive Event Shape to Veto Jets,* *Phys.Rev.Lett.* **105** (2010) 092002, [arXiv:1004.2489].

[105] C. W. Bauer, F. J. Tackmann, J. R. Walsh, and S. Zuberi, *Factorization and Resummation for Dijet Invariant Mass Spectra,* *Phys.Rev.* **D85** (2012) 074006, [arXiv:1106.6047].

[106] G. Soyez, G. P. Salam, J. Kim, S. Dutta, and M. Cacciari, *Pileup subtraction for jet shapes,* *Phys.Rev.Lett.* **110** (2013), no. 16 162001, [arXiv:1211.2811].

[107] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,* arXiv:1405.0301.

[108] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual,* *JHEP* **0605** (2006) 026, [hep-ph/0603175].

[109] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1,* *Comput.Phys.Commun.* **178** (2008) 852–867, [arXiv:0710.3820].

[110] G. Marchesini, B. Webber, G. Abbiendi, I. Knowles, M. Seymour, et al., *HERWIG: A Monte Carlo event generator for simulating hadron emission reactions with interfering gluons. Version 5.1 - April 1991,* *Comput.Phys.Commun.* **67** (1992) 465–508.

[111] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes),* *JHEP* **0101** (2001) 010, [hep-ph/0011363].

[112] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al., *HERWIG 6.5 release note,* hep-ph/0210213.

[113] M. Bahr, S. Gieseke, M. Gigg, D. Grellscheid, K. Hamilton, et al., *Herwig++ Physics and Manual,* *Eur.Phys.J.* **C58** (2008) 639–707, [arXiv:0803.0883].

[114] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, Phys. Rev. Lett. **100** (2008) 242001, [arXiv:0802.2470].

[115] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, *Techniques for improved heavy particle searches with jet substructure*, Phys.Rev. **D80** (2009) 051501, [arXiv:0903.5081].

[116] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, *Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches*, Phys.Rev. **D81** (2010) 094023, [arXiv:0912.0033].

[117] D. Krohn, J. Thaler, and L.-T. Wang, *Jet Trimming*, JHEP **1002** (2010) 084, [arXiv:0912.1342].

[118] D. Krohn, M. D. Schwartz, M. Low, and L.-T. Wang, *Jet Cleansing: Pileup Removal at High Luminosity*, arXiv:1309.4777.

[119] P. Berta, M. Spousta, D. W. Miller, and R. Leitner, *Particle-level pileup subtraction for jets and jet shapes*, JHEP **1406** (2014) 092, [arXiv:1403.3108].

[120] M. Cacciari, G. P. Salam, and G. Soyez, *SoftKiller, a particle-level pileup removal method*, arXiv:1407.0408.

[121] D. Bertolini, P. Harris, M. Low, and N. Tran, *Pileup Per Particle Identification*, arXiv:1407.6013.

[122] J. Gallicchio and M. D. Schwartz, *Quark and Gluon Tagging at the LHC*, Phys.Rev.Lett. **107** (2011) 172001, [arXiv:1106.3076].

[123] J. Gallicchio and M. D. Schwartz, *Quark and Gluon Jet Substructure*, JHEP **1304** (2013) 090, [arXiv:1211.7038].

[124] J. Gallicchio and M. D. Schwartz, *Pure Samples of Quark and Gluon Jets at the LHC*, JHEP **1110** (2011) 103, [arXiv:1104.1175].

[125] D. Adams, A. Arce, L. Asquith, M. Backovic, T. Barillari, et al., *Towards an Understanding of the Correlations in Jet Substructure*, arXiv:1504.00679.

[126] **CMS** Collaboration, *V Tagging Observables and Correlations*, Tech. Rep. CMS-PAS-JME-14-002, 2014.

[127] **ATLAS** Collaboration, G. Aad et al., *Measurement of the cross-section of high transverse momentum vector bosons reconstructed as single jets and studies of jet substructure in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, New J.Phys. **16** (2014), no. 11 113013, [arXiv:1407.0800].

[128] **CMS** Collaboration, V. Khachatryan et al., *Search for vector-like T quarks decaying to top quarks and Higgs bosons in the all-hadronic channel using jet substructure*, arXiv:1503.01952.

[129] **CMS** Collaboration, C. Collaboration, *Search for pair-produced vector-like top quark partners decaying to bW in the fully hadronic channel using jet substructure at 8 TeV*, .

[130] **CMS** Collaboration, V. Khachatryan et al., *Search for a massive resonance decaying into a Higgs boson and a W or Z boson in hadronic final states in proton-proton collisions at sqrt(s) = 8 TeV*, arXiv:1506.01443.

[131] **ATLAS** Collaboration, G. Aad et al., *Search for high-mass diboson resonances with boson-tagged jets in proton-proton collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector*, arXiv:1506.00962.

[132] M. Field, G. Gur-Ari, D. A. Kosower, L. Mannelli, and G. Perez, *Three-Prong Distribution of Massive Narrow QCD Jets*, Phys.Rev. **D87** (2013), no. 9 094013, [arXiv:1212.2106].

[133] M. Dasgupta, A. Powling, and A. Siodmok, *On jet substructure methods for signal jets*, JHEP **08** (2015) 079, [arXiv:1503.01088].

[134] M. Seymour, *Jet shapes in hadron collisions: Higher orders, resummation and hadronization*, Nucl.Phys. **B513** (1998) 269–300, [hep-ph/9707338].

[135] H.-n. Li, Z. Li, and C.-P. Yuan, *QCD resummation for jet substructures*, Phys. Rev. Lett. **107** (2011) 152001, [arXiv:1107.4535].

[136] A. J. Larkoski, *QCD Analysis of the Scale-Invariance of Jets*, Phys. Rev. D **86** (2012) 054004, [arXiv:1207.1437].

[137] M. Jankowiak and A. J. Larkoski, *Angular Scaling in Jets*, JHEP **1204** (2012) 039, [arXiv:1201.2688].

[138] Y.-T. Chien and I. Vitev, *Jet Shape Resummation Using Soft-Collinear Effective Theory*, arXiv:1405.4293.

[139] Y.-T. Chien, *Resummation of Jet Shapes and Extracting Properties of the Quark-Gluon Plasma*, Int.J.Mod.Phys.Conf.Ser. **37** (2015) 1560047, [arXiv:1411.0741].

[140] J. Isaacson, H.-n. Li, Z. Li, and C. P. Yuan, *Factorization for substructures of boosted Higgs jets*, arXiv:1505.06368.

[141] D. Krohn, M. D. Schwartz, T. Lin, and W. J. Waalewijn, *Jet Charge at the LHC*, Phys. Rev. Lett. **110** (2013) 212001, [arXiv:1209.2421].

476

[142] W. J. Waalewijn, *Calculating the Charge of a Jet*, *Phys.Rev.* **D86** (2012) 094030, [arXiv:1209.3019].

[143] D. Bertolini, J. Thaler, and J. R. Walsh, *The First Calculation of Fractional Jets*, *JHEP* **05** (2015) 008, [arXiv:1501.01965].

[144] B. Bhattacherjee, S. Mukhopadhyay, M. M. Nojiri, Y. Sakaki, and B. R. Webber, *Associated jet and subjet rates in light-quark and gluon jet discrimination*, *JHEP* **1504** (2015) 131, [arXiv:1501.04794].

[145] L. G. Almeida, S. J. Lee, G. Perez, G. F. Sterman, I. Sung, et al., *Substructure of high-p_T Jets at the LHC*, *Phys.Rev.* **D79** (2009) 074017, [arXiv:0807.0234].

[146] A. J. Larkoski, S. Marzani, and J. Thaler, *Sudakov Safety in Perturbative QCD*, *Phys.Rev.* **D91** (2015), no. 11 111501, [arXiv:1502.01719].

[147] G. P. Korchemsky and G. Sterman, *Power corrections to event shapes and factorization*, *Nucl. Phys.* **B 555** (1999) 335–351, [hep-ph/9902341].

[148] G. P. Korchemsky and S. Tafat, *On power corrections to the event shape distributions in QCD*, *JHEP* **10** (2000) 010, [hep-ph/0007005].

[149] W. T. Giele, D. A. Kosower, and P. Z. Skands, *A simple shower and matching algorithm*, *Phys. Rev.* **D78** (2008) 014026, [arXiv:0707.3652].

[150] W. T. Giele, D. A. Kosower, and P. Z. Skands, *Higher-Order Corrections to Timelike Jets*, *Phys. Rev.* **D84** (2011) 054003, [arXiv:1102.2126].

[151] A. Gehrmann-De Ridder, M. Ritzmann, and P. Z. Skands, *Timelike Dipole-Antenna Showers with Massive Fermions*, *Phys. Rev.* **D85** (2012) 014013, [arXiv:1108.6172].

[152] M. Ritzmann, D. A. Kosower, and P. Skands, *Antenna Showers with Hadronic Initial States*, *Phys. Lett.* **B718** (2013) 1345–1350, [arXiv:1210.6345].

[153] L. Hartgring, E. Laenen, and P. Skands, *Antenna Showers with One-Loop Matrix Elements*, *JHEP* **10** (2013) 127, [arXiv:1303.4974].

[154] A. J. Larkoski, J. J. Lopez-Villarejo, and P. Skands, *Helicity-Dependent Showers and Matching with VINCIA*, *Phys. Rev.* **D87** (2013), no. 5 054033, [arXiv:1301.0933].

[155] G. Salam and D. Wicke, *Hadron masses and power corrections to event shapes*, *JHEP* **0105** (2001) 061, [hep-ph/0102343].

[156] V. Mateu, I. W. Stewart, and J. Thaler, *Power Corrections to Event Shapes with Mass-Dependent Operators*, *Phys.Rev.* **D87** (2013), no. 1 014025, [arXiv:1209.3781].

[157] S. Brandt and H. Dahmen, *Axes and Scalar Measures of Two-Jet and Three-Jet Events*, *Z.Phys.* **C1** (1979) 61.

477

[158] S. Catani, Y. L. Dokshitzer, M. Seymour, and B. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, Nucl.Phys. **B406** (1993) 187–224.

[159] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys.Rev. **D48** (1993) 3160–3166, [hep-ph/9305266].

[160] Y. L. Dokshitzer, G. Leder, S. Moretti, and B. Webber, *Better jet clustering algorithms*, JHEP **9708** (1997) 001, [hep-ph/9707323].

[161] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, hep-ph/9907280.

[162] M. Wobisch, *Measurement and QCD analysis of jet cross-sections in deep inelastic positron proton collisions at $\sqrt{s} = 300$ GeV*, 2000.

[163] R. Appleby and M. Seymour, *Nonglobal logarithms in interjet energy flow with kt clustering requirement*, JHEP **0212** (2002) 063, [hep-ph/0211426].

[164] A. Banfi and M. Dasgupta, *Problems in resumming interjet energy flows with k_t clustering*, Phys. Lett. B **628** (2005) 49–56, [hep-ph/0508159].

[165] A. Banfi, M. Dasgupta, K. Khelifa-Kerfa, and S. Marzani, *Non-global logarithms and jet algorithms in high-pT jet shapes*, JHEP **08** (2010) 064, [arXiv:1004.3483].

[166] R. Kelley, J. R. Walsh, and S. Zuberi, *Abelian Non-Global Logarithms from Soft Gluon Clustering*, JHEP **1209** (2012) 117, [arXiv:1202.2361].

[167] S. D. Ellis, C. K. Vermilion, J. R. Walsh, A. Hornig, and C. Lee, *Jet Shapes and Jet Algorithms in SCET*, JHEP **1011** (2010) 101, [arXiv:1001.0014].

[168] S. Fleming, A. H. Hoang, S. Mantry, and I. W. Stewart, *Top Jets in the Peak Region: Factorization Analysis with NLL Resummation*, Phys. Rev. D **77** (2008) 114003, [arXiv:0711.2079].

[169] R. Kelley, M. D. Schwartz, R. M. Schabinger, and H. X. Zhu, *Jet Mass with a Jet Veto at Two Loops and the Universality of Non-Global Structure*, Phys.Rev. **D86** (2012) 054017, [arXiv:1112.3343].

[170] A. Hornig, C. Lee, J. R. Walsh, and S. Zuberi, *Double Non-Global Logarithms In-N-Out of Jets*, JHEP **1201** (2012) 149, [arXiv:1110.0004].

[171] Z. Nagy and Z. Trocsanyi, *Next-to-leading order calculation of four jet shape variables*, Phys.Rev.Lett. **79** (1997) 3604–3607, [hep-ph/9707309].

[172] Z. Nagy and Z. Trocsanyi, *Next-to-leading order calculation of four jet observables in electron positron annihilation*, Phys.Rev. **D59** (1999) 014020, [hep-ph/9806317].

[173] Z. Nagy, *Three jet cross-sections in hadron hadron collisions at next-to-leading order*, *Phys.Rev.Lett.* **88** (2002) 122003, [hep-ph/0110315].

[174] Z. Nagy and Z. Trocsanyi, *Multijet cross-sections in deep inelastic scattering at next-to-leading order*, *Phys.Rev.Lett.* **87** (2001) 082001, [hep-ph/0104315].

[175] Z. Nagy, *Next-to-leading order calculation of three jet observables in hadron hadron collision*, *Phys. Rev. D* **68** (2003) 094002, [hep-ph/0307268].

[176] A. V. Manohar and I. W. Stewart, *The Zero-Bin and Mode Factorization in Quantum Field Theory*, *Phys.Rev.* **D76** (2007) 074002, [hep-ph/0605001].

[177] N. Fischer, S. Gieseke, S. Plätzer, and P. Skands, *Revisiting radiation patterns in e^+e^- collisions*, *Eur.Phys.J.* **C74** (2014), no. 4 2831, [arXiv:1402.3186].

[178] **OPAL** Collaboration, N. Fischer, S. Gieseke, S. Kluth, S. Plätzer, and P. Skands, *Measurement of observables sensitive to coherence effects in hadronic Z decays with the OPAL detector at LEP*, arXiv:1505.01636.

[179] A. Buckley, J. Butterworth, S. Gieseke, D. Grellscheid, S. Hoche, et al., *General-purpose event generators for LHC physics*, *Phys. Rept.* **504** (2011) 145–233, [arXiv:1101.2599].

[180] P. Skands, *Introduction to QCD*, arXiv:1207.2389.

[181] M. H. Seymour and M. Marx, *Monte Carlo Event Generators*, arXiv:1304.6677.

[182] S. Gieseke, *Simulation of jets at colliders*, *Prog.Part.Nucl.Phys.* **72** (2013) 155–205.

[183] S. Höche, *Introduction to parton-shower event generators*, arXiv:1411.4085.

[184] T. Gleisberg, S. Hoeche, F. Krauss, A. Schalicke, S. Schumann, et al., *SHERPA 1. alpha: A Proof of concept version*, *JHEP* **0402** (2004) 056, [hep-ph/0311263].

[185] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al., *Event generation with SHERPA 1.1*, *JHEP* **0902** (2009) 007, [arXiv:0811.4622].

[186] L. Lonnblad, *ARIADNE version 4: A Program for simulation of QCD cascades implementing the color dipole model*, *Comput.Phys.Commun.* **71** (1992) 15–31.

[187] S. Höche and S. Prestel, *The midpoint between dipole and parton showers*, arXiv:1506.05057.

[188] S. Platzer and S. Gieseke, *Dipole Showers and Automated NLO Matching in Herwig++*, *Eur.Phys.J.* **C72** (2012) 2187, [arXiv:1109.6256].

[189] S. Platzer and S. Gieseke, *Coherent Parton Showers with Local Recoils*, *JHEP* **1101** (2011) 024, [arXiv:0909.5593].

[190] M. Dasgupta and G. P. Salam, *Accounting for coherence in interjet E(t) flow: A Case study*, *JHEP* **0203** (2002) 017, [hep-ph/0203009].

[191] F. J. Tackmann, J. R. Walsh, and S. Zuberi, *Resummation Properties of Jet Vetoes at the LHC*, *Phys.Rev.* **D86** (2012) 053011, [arXiv:1206.4312].

[192] M. Dasgupta, F. Dreyer, G. P. Salam, and G. Soyez, *Small-radius jets to all orders in QCD*, *JHEP* **1504** (2015) 039, [arXiv:1411.5182].

[193] S. Bosch, B. Lange, M. Neubert, and G. Paz, *Factorization and shape function effects in inclusive B meson decays*, *Nucl.Phys.* **B699** (2004) 335–386, [hep-ph/0402094].

[194] A. H. Hoang and I. W. Stewart, *Designing Gapped Soft Functions for Jet Production*, *Phys. Lett. B* **660** (2008) 483–493, [arXiv:0709.3519].

[195] Z. Ligeti, I. W. Stewart, and F. J. Tackmann, *Treating the b quark distribution function with reliable uncertainties*, *Phys. Rev. D* **78** (2008) 114014, [arXiv:0807.1926].

[196] R. Akhoury and V. I. Zakharov, *On the universality of the leading, 1/Q power corrections in QCD*, *Phys.Lett.* **B357** (1995) 646–652, [hep-ph/9504248].

[197] Y. L. Dokshitzer and B. R. Webber, *Calculation of power corrections to hadronic event shapes*, *Phys. Lett. B* **352** (1995) 451–455, [hep-ph/9504219].

[198] C. Lee and G. F. Sterman, *Universality of nonperturbative effects in event shapes*, *eConf* **C0601121** (2006) A001, [hep-ph/0603066].

[199] C. Lee, *Universal nonperturbative effects in event shapes from soft-collinear effective theory*, *Mod.Phys.Lett.* **A22** (2007) 835–851, [hep-ph/0703030].

[200] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *Dissecting Soft Radiation with Factorization*, *Phys.Rev.Lett.* **114** (2015), no. 9 092001, [arXiv:1405.6722].

[201] M. Beneke, *Renormalons*, *Phys.Rept.* **317** (1999) 1–142, [hep-ph/9807443].

[202] E. Gardi, *Perturbative and nonperturbative aspects of moments of the thrust distribution in e+ e- annihilation*, *JHEP* **0004** (2000) 030, [hep-ph/0003179].

[203] A. Hornig, C. Lee, and G. Ovanesyan, *Effective Predictions of Event Shapes: Factorized, Resummed, and Gapped Angularity Distributions*, *JHEP* **0905** (2009) 122, [arXiv:0901.3780].

[204] **L3** Collaboration, P. Achard et al., *Studies of hadronic event structure in e^+e^- annihilation from 30-GeV to 209-GeV with the L3 detector*, *Phys.Rept.* **399** (2004) 71–174, [hep-ex/0406049].

[205] T. Gehrmann, M. Jaquier, and G. Luisoni, *Hadronization effects in event shape moments, Eur.Phys.J.* **C67** (2010) 57–72, [arXiv:0911.2422].

[206] R. Abbate, M. Fickinger, A. H. Hoang, V. Mateu, and I. W. Stewart, *Precision Thrust Cumulant Moments at N^3LL, Phys. Rev.* **D 86** (2012) 094002, [arXiv:1204.5746].

[207] A. H. Hoang, D. W. Kolodrubetz, V. Mateu, and I. W. Stewart, *A Precise Determination of α_s from the C-parameter Distribution*, arXiv:1501.04111.

[208] X. Liu and F. Petriello, *Resummation of jet-veto logarithms in hadronic processes containing jets, Phys. Rev.* **D 87** (2013) 014018, [arXiv:1210.1906].

[209] X. Liu and F. Petriello, *Reducing theoretical uncertainties for exclusive Higgs-boson plus one-jet production at the LHC, Phys.Rev.* **D87** (2013), no. 9 094027, [arXiv:1303.4405].

[210] J. Talbert, *Automated Calculations of Dijet Soft Functions, SCET Conference* (2015).

[211] R. Boughezal, X. Liu, and F. Petriello, *N-jettiness soft function at next-to-next-to-leading order, Phys.Rev.* **D91** (2015), no. 9 094035, [arXiv:1504.02540].

[212] A. Banfi, G. Marchesini, and G. Smye, *Away-from-jet energy flow, JHEP* **08** (2002) 006, [hep-ph/0206076].

[213] H. Weigert, *Nonglobal jet evolution at finite N(c), Nucl.Phys.* **B685** (2004) 321–350, [hep-ph/0312050].

[214] Y. Hatta and T. Ueda, *Resummation of non-global logarithms at finite N_c, Nucl. Phys.* **B874** (2013) 808–820, [arXiv:1304.6930].

[215] S. Caron-Huot, *Resummation of non-global logarithms and the BFKL equation*, arXiv:1501.03754.

[216] R. Boughezal, C. Focke, W. Giele, X. Liu, and F. Petriello, *Higgs boson production in association with a jet using jettiness subtraction, Phys.Lett.* **B748** (2015) 5–8, [arXiv:1505.03893].

[217] R. Boughezal, C. Focke, X. Liu, and F. Petriello, *W-boson production in association with a jet at next-to-next-to-leading order in perturbative QCD*, arXiv:1504.02131.

[218] J. Gaunt, M. Stahlhofen, F. J. Tackmann, and J. R. Walsh, *N-jettiness Subtractions for NNLO QCD Calculations*, arXiv:1505.04794.

[219] R. Boughezal, F. Caola, K. Melnikov, F. Petriello, and M. Schulze, *Higgs Boson Production in Association with a Jet at Next-to-Next-to-Leading Order*, arXiv:1504.07922.

481

[220] A. J. Larkoski, F. Maltoni, and M. Selvaggi, *Tracking down hyper-boosted top quarks*, *JHEP* **1506** (2015) 032, [arXiv:1503.03347].

[221] S. Fleming, A. H. Hoang, S. Mantry, and I. W. Stewart, *Jets from massive unstable particles: Top-mass determination*, *Phys. Rev. D* **77** (2008) 074010, [hep-ph/0703207].

[222] P. De Causmaecker, R. Gastmans, W. Troost, and T. T. Wu, *Multiple Bremsstrahlung in Gauge Theories at High-Energies. 1: General Formalism for Quantum Electrodynamics*, *Nucl. Phys. B* **206** (1982) 53.

[223] F. A. Berends, R. Kleiss, P. De Causmaecker, R. Gastmans, W. Troost, et al., *Multiple Bremsstrahlung in Gauge Theories at High-Energies. 2. Single Bremsstrahlung*, *Nucl. Phys. B* **206** (1982) 61.

[224] J. Gunion and Z. Kunszt, *Improved Analytic Techniques for Tree Graph Calculations and the G g q anti-q Lepton anti-Lepton Subprocess*, *Phys. Lett. B* **161** (1985) 333.

[225] Z. Xu, D.-H. Zhang, and L. Chang, *Helicity Amplitudes for Multiple Bremsstrahlung in Massless Nonabelian Gauge Theories*, *Nucl. Phys. B* **291** (1987) 392.

[226] F. A. Berends and W. Giele, *Recursive Calculations for Processes with n Gluons*, *Nucl.Phys.* **B306** (1988) 759.

[227] M. L. Mangano, S. J. Parke, and Z. Xu, *Duality and Multi - Gluon Scattering*, *Nucl.Phys.* **B298** (1988) 653.

[228] M. L. Mangano, *The Color Structure of Gluon Emission*, *Nucl.Phys.* **B309** (1988) 461.

[229] Z. Bern and D. A. Kosower, *Color decomposition of one loop amplitudes in gauge theories*, *Nucl.Phys.* **B362** (1991) 389–448.

[230] Z. Bern, L. J. Dixon, D. C. Dunbar, and D. A. Kosower, *One-Loop n-Point Gauge Theory Amplitudes, Unitarity and Collinear Limits*, *Nucl. Phys. B* **425** (1994) 217–260, [hep-ph/9403226].

[231] Z. Bern, L. J. Dixon, D. C. Dunbar, and D. A. Kosower, *Fusing gauge theory tree amplitudes into loop amplitudes*, *Nucl.Phys.* **B435** (1995) 59–101, [hep-ph/9409265].

[232] W. Giele and E. Glover, *Higher order corrections to jet cross-sections in e+ e- annihilation*, *Phys. Rev. D* **46** (1992) 1980–2010.

[233] P. B. Arnold and M. Reno, *The Complete Computation of High p(t) W and Z Production in 2nd Order QCD*, *Nucl. Phys. B* **319** (1989) 37.

[234] W. Giele, E. Glover, and D. A. Kosower, *Higher order corrections to jet cross-sections in hadron colliders*, *Nucl. Phys. B* **403** (1993) 633–670, [hep-ph/9302225].

482

[235] Z. Bern, L. J. Dixon, and D. A. Kosower, *One-loop amplitudes for e+ e- to four partons*, Nucl. Phys. B **513** (1998) 3–86, [hep-ph/9708239].

[236] R. Ellis, W. Giele, Z. Kunszt, K. Melnikov, and G. Zanderighi, *One-loop amplitudes for W+3 jet production in hadron collisions*, JHEP **01** (2009) 012, [arXiv:0810.2762].

[237] C. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, et al., *Precise Predictions for W + 3 Jet Production at Hadron Colliders*, Phys.Rev.Lett. **102** (2009) 222001, [arXiv:0902.2760].

[238] C. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, et al., *Next-to-Leading Order QCD Predictions for W+3-Jet Distributions at Hadron Colliders*, Phys.Rev. **D80** (2009) 074036, [arXiv:0907.1984].

[239] C. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, et al., *Next-to-Leading Order QCD Predictions for Z, gamma* +3-Jet Distributions at the Tevatron*, Phys. Rev. D **82** (2010) 074002, [arXiv:1004.1659].

[240] C. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, et al., *Precise Predictions for W + 4 Jet Production at the Large Hadron Collider*, Phys.Rev.Lett. **106** (2011) 092001, [arXiv:1009.2338].

[241] H. Ita, Z. Bern, L. Dixon, F. Febres Cordero, D. Kosower, et al., *Precise Predictions for Z + 4 Jets at Hadron Colliders*, Phys.Rev. **D85** (2012) 031501, [arXiv:1108.2229].

[242] Z. Bern, L. Dixon, F. Febres Cordero, S. Höche, H. Ita, et al., *Next-to-Leading Order W + 5-Jet Production at the LHC*, Phys.Rev. **D88** (2013), no. 1 014025, [arXiv:1304.1253].

[243] Z. Bern, L. J. Dixon, and D. A. Kosower, *One loop corrections to five gluon amplitudes*, Phys. Rev. Lett. **70** (1993) 2677–2680, [hep-ph/9302280].

[244] Z. Kunszt, A. Signer, and Z. Trocsanyi, *One loop helicity amplitudes for all 2 → 2 processes in QCD and N=1 supersymmetric Yang-Mills theory*, Nucl. Phys. B **411** (1994) 397–442, [hep-ph/9305239].

[245] Z. Kunszt, A. Signer, and Z. Trocsanyi, *One loop radiative corrections to the helicity amplitudes of QCD processes involving four quarks and one gluon*, Phys. Lett. B **336** (1994) 529–536, [hep-ph/9405386].

[246] Z. Bern, L. J. Dixon, and D. A. Kosower, *One loop corrections to two quark three gluon amplitudes*, Nucl. Phys. B **437** (1995) 259–304, [hep-ph/9409393].

[247] Z. Bern, G. Diana, L. Dixon, F. Febres Cordero, S. Hoeche, et al., *Four-Jet Production at the Large Hadron Collider at Next-to-Leading Order in QCD*, *Phys.Rev.Lett.* **109** (2012) 042001, [arXiv:1112.3940].

[248] S. Badger, B. Biedermann, P. Uwer, and V. Yundin, *NLO QCD corrections to multi-jet production at the LHC with a centre-of-mass energy of* $\sqrt{s} = 8$ *TeV*, *Phys. Lett.* **B718** (2013) 965–978, [arXiv:1209.0098].

[249] S. Badger, B. Biedermann, P. Uwer, and V. Yundin, *Next-to-leading order QCD corrections to five jet production at the LHC*, *Phys. Rev. D* **89** (2014) 034019, [arXiv:1309.6585].

[250] J. M. Campbell, R. K. Ellis, and G. Zanderighi, *Next-to-Leading order Higgs + 2 jet production via gluon fusion*, *JHEP* **10** (2006) 028, [hep-ph/0608194].

[251] R. P. Kauffman, S. V. Desai, and D. Risal, *Production of a Higgs boson plus two jets in hadronic collisions*, *Phys. Rev. D* **55** (1997) 4005–4015, [hep-ph/9610541].

[252] C. R. Schmidt, $H \rightarrow g\,g\,g\ (g\,q\,anti\text{-}q)$ *at two loops in the large-M(t) limit*, *Phys. Lett. B* **413** (1997) 391–395, [hep-ph/9707448].

[253] L. J. Dixon and Y. Sofianatos, *Analytic one-loop amplitudes for a Higgs boson plus four partons*, *JHEP* **0908** (2009) 058, [arXiv:0906.0008].

[254] S. Badger, E. Nigel Glover, P. Mastrolia, and C. Williams, *One-loop Higgs plus four gluon amplitudes: Full analytic results*, *JHEP* **1001** (2010) 036, [arXiv:0909.4475].

[255] J. M. Campbell, R. K. Ellis, and C. Williams, *Hadronic production of a Higgs boson and two jets at next-to-leading order*, *Phys. Rev. D* **81** (2010) 074023, [arXiv:1001.4495].

[256] H. van Deurzen, N. Greiner, G. Luisoni, P. Mastrolia, E. Mirabella, et al., *NLO QCD corrections to the production of Higgs plus two jets at the LHC*, *Phys. Lett. B* **721** (2013) 74–81, [arXiv:1301.0493].

[257] G. Cullen, H. van Deurzen, N. Greiner, G. Luisoni, P. Mastrolia, et al., *Next-to-Leading-Order QCD Corrections to Higgs Boson Production Plus Three Jets in Gluon Fusion*, *Phys. Rev. Lett.* **111** (2013), no. 13 131801, [arXiv:1307.4737].

[258] F. Campanario, T. M. Figy, S. Plätzer, and M. Sjödahl, *Electroweak Higgs Boson Plus Three Jet Production at Next-to-Leading-Order QCD*, *Phys. Rev. Lett.* **111** (2013), no. 21 211802, [arXiv:1308.2932].

[259] **SM and NLO Multileg Working Group** Collaboration, J. R. Andersen et al., *The SM and NLO multileg working group: Summary report*, arXiv:1003.1241.

[260] **SM and NLO MULTILEG Working Group, SM MC Working Group** Collaboration, J. Alcaraz Maestre et al., *The SM and NLO Multileg and SM MC Working Groups: Summary Report*, arXiv:1203.6803.

[261] G. Ossola, C. G. Papadopoulos, and R. Pittau, *CutTools: a program implementing the OPP reduction method to compute one-loop amplitudes*, JHEP **03** (2008) 042, [arXiv:0711.3596].

[262] C. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, et al., *An Automated Implementation of On-Shell Methods for One-Loop Amplitudes*, Phys.Rev. D **78** (2008) 036003, [arXiv:0803.4180].

[263] T. Binoth, J.-P. Guillet, G. Heinrich, E. Pilon, and T. Reiter, *Golem95: A Numerical program to calculate one-loop tensor integrals with up to six external legs*, Comput. Phys. Commun. **180** (2009) 2317–2330, [arXiv:0810.0992].

[264] P. Mastrolia, G. Ossola, T. Reiter, and F. Tramontano, *Scattering AMplitudes from Unitarity-based Reduction Algorithm at the Integrand-level*, JHEP **08** (2010) 080, [arXiv:1006.0710].

[265] S. Badger, B. Biedermann, and P. Uwer, *NGluon: A Package to Calculate One-loop Multi-gluon Amplitudes*, Comput. Phys. Commun. **182** (2011) 1674–1692, [arXiv:1011.2900].

[266] J. Fleischer and T. Riemann, *A Complete algebraic reduction of one-loop tensor Feynman integrals*, Phys. Rev. D **83** (2011) 073004, [arXiv:1009.4436].

[267] G. Cullen, J.-P. Guillet, G. Heinrich, T. Kleinschmidt, E. Pilon, et al., *Golem95C: A library for one-loop integrals with complex masses*, arXiv:1101.5595.

[268] V. Hirschi et al., *Automation of one-loop QCD corrections*, JHEP **05** (2011) 044, [arXiv:1103.0621].

[269] G. Bevilacqua, M. Czakon, M. V. Garzelli, A. van Hameren, A. Kardos, C. G. Papadopoulos, R. Pittau, and M. Worek, *HELAC-NLO*, Comput. Phys. Commun. **184** (2013) 986–997, [arXiv:1110.1499].

[270] G. Cullen, N. Greiner, G. Heinrich, G. Luisoni, P. Mastrolia, et al., *Automated One-Loop Calculations with GoSam*, Eur. Phys. J. **C72** (2012) 1889, [arXiv:1111.2034].

[271] F. Cascioli, P. Maierhofer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, [arXiv:1111.5206].

[272] S. Badger, B. Biedermann, P. Uwer, and V. Yundin, *Numerical evaluation of virtual corrections to multi-jet production in massless QCD*, Comput. Phys. Commun. **184** (2013) 1981–1998, [arXiv:1209.0100].

[273] S. Actis, A. Denner, L. Hofer, A. Scharf, and S. Uccirati, *Recursive generation of one-loop amplitudes in the Standard Model*, *JHEP* **1304** (2013) 037, [arXiv:1211.6316].

[274] T. Peraro, *Ninja: Automated Integrand Reduction via Laurent Expansion for One-Loop Amplitudes*, *Comput. Phys. Commun.* **185** (2014) 2771–2797, [arXiv:1403.1229].

[275] G. Cullen, H. van Deurzen, N. Greiner, G. Heinrich, G. Luisoni, et al., *GOSAM-2.0: a tool for automated one-loop calculations within the Standard Model and beyond*, *Eur. Phys. J.* **C74** (2014), no. 8 3001, [arXiv:1404.7096].

[276] T. Becher and M. D. Schwartz, *A Precise determination of α_s from LEP thrust data using effective field theory*, *JHEP* **07** (2008) 034, [arXiv:0803.0342].

[277] S. Alioli, C. W. Bauer, C. J. Berggren, A. Hornig, F. J. Tackmann, et al., *Combining Higher-Order Resummation with Multiple NLO Calculations and Parton Showers in GENEVA*, *JHEP* **1309** (2013) 120, [arXiv:1211.7049].

[278] S. Alioli, C. W. Bauer, C. Berggren, F. J. Tackmann, J. R. Walsh, and S. Zuberi, *Matching Fully Differential NNLO Calculations and Parton Showers*, *JHEP* **06** (2014) 089, [arXiv:1311.0286].

[279] S. Alioli, C. W. Bauer, C. Berggren, F. J. Tackmann, and J. R. Walsh, *Drell-Yan production at NNLL'+NNLO matched to parton showers*, *Phys. Rev. D* **92** (2015), no. 9 094020, [arXiv:1508.01475].

[280] P. A. Baikov, K. G. Chetyrkin, A. V. Smirnov, V. A. Smirnov, and M. Steinhauser, *Quark and gluon form factors to three loops*, *Phys. Rev. Lett.* **102** (2009) 212002, [arXiv:0902.3519].

[281] T. Gehrmann, E. W. N. Glover, T. Huber, N. Ikizlerli, and C. Studerus, *Calculation of the quark and gluon form factors to three loops in QCD*, arXiv:1004.3653.

[282] T. Becher and M. D. Schwartz, *Direct photon production with effective field theory*, *JHEP* **02** (2010) 040, [arXiv:0911.0681].

[283] T. Becher, C. Lorentzen, and M. D. Schwartz, *Resummation for W and Z production at large pT*, *Phys. Rev. Lett.* **108** (2012) 012001, [arXiv:1106.4310].

[284] T. Becher, C. Lorentzen, and M. D. Schwartz, *Precision Direct Photon and W-Boson Spectra at High p_T and Comparison to LHC Data*, *Phys. Rev. D* **86** (2012) 054026, [arXiv:1206.6115].

[285] T. Becher, G. Bell, C. Lorentzen, and S. Marti, *Transverse-momentum spectra of electroweak bosons near threshold at NNLO*, *JHEP* **1402** (2014) 004, [arXiv:1309.3245].

[286] F. P. Huang, C. S. Li, H. T. Li, and J. Wang, *Renormalization-group improved predictions for Higgs boson production at large p_T*, Phys. Rev. D **90** (2014), no. 9 094024, [arXiv:1406.2591].

[287] T. Becher, G. Bell, C. Lorentzen, and S. Marti, *The transverse-momentum spectrum of Higgs bosons near threshold at NNLO*, JHEP **1411** (2014) 026, [arXiv:1407.4111].

[288] J.-y. Chiu, R. Kelley, and A. V. Manohar, *Electroweak Corrections using Effective Field Theory: Applications to the LHC*, Phys. Rev. D **78** (2008) 073006, [arXiv:0806.1240].

[289] H. X. Zhu, C. S. Li, J. Wang, and J. J. Zhang, *Factorization and resummation of s-channel single top quark production*, JHEP **1102** (2011) 099, [arXiv:1006.0681].

[290] R. Kelley and M. D. Schwartz, *Threshold Hadronic Event Shapes with Effective Field Theory*, Phys. Rev. D **83** (2011) 033001, [arXiv:1008.4355].

[291] J. Wang, C. S. Li, H. X. Zhu, and J. J. Zhang, *Factorization and resummation of t-channel single top quark production*, arXiv:1010.4509.

[292] V. Ahrens, A. Ferroglia, M. Neubert, B. D. Pecjak, and L. L. Yang, *Renormalization-Group Improved Predictions for Top-Quark Pair Production at Hadron Colliders*, JHEP **09** (2010) 097, [arXiv:1003.5827].

[293] V. Ahrens, A. Ferroglia, M. Neubert, B. D. Pecjak, and L.-L. Yang, *RG-improved single-particle inclusive cross sections and forward-backward asymmetry in $t\bar{t}$ production at hadron colliders*, JHEP **1109** (2011) 070, [arXiv:1103.0550].

[294] H. T. Li, C. S. Li, and S. A. Li, *Renormalization group improved predictions for $t\bar{t}W^{\pm}$ production at hadron colliders*, Phys. Rev. D **90** (2014), no. 9 094009, [arXiv:1409.1460].

[295] Z. L. Liu, C. S. Li, J. Wang, and Y. Wang, *Resummation prediction on the jet mass spectrum in one-jet inclusive production at the LHC*, JHEP **04** (2015) 005, [arXiv:1412.1337].

[296] R. Kelley and M. D. Schwartz, *1-loop matching and NNLL resummation for all partonic 2 to 2 processes in QCD*, Phys. Rev. D **83** (2011) 045022, [arXiv:1008.2759].

[297] A. Broggio, A. Ferroglia, B. D. Pecjak, and Z. Zhang, *NNLO hard functions in massless QCD*, JHEP **12** (2014) 005, [arXiv:1409.5294].

[298] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *Combining Fixed-Order Helicity Amplitudes With Resummation Using SCET*, PoS **LL2012** (2012) 058, [arXiv:1211.2305].

487

[299] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *The Beam Thrust Cross Section for Drell-Yan at NNLL Order*, Phys. Rev. Lett. **106** (2011) 032001, [arXiv:1005.4060].

[300] C. F. Berger, C. Marcantonini, I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *Higgs Production with a Central Jet Veto at NNLL+NNLO*, JHEP **1104** (2011) 092, [arXiv:1012.4480].

[301] T. T. Jouttenus, I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *The Soft Function for Exclusive N-Jet Production at Hadron Colliders*, Phys.Rev. **D83** (2011) 114030, [arXiv:1102.4344].

[302] D. Kang, C. Lee, and I. W. Stewart, *Using 1-Jettiness to Measure 2 Jets in DIS 3 Ways*, Phys. Rev. **D88** (2013) 054004, [arXiv:1303.6952].

[303] Z.-B. Kang, X. Liu, and S. Mantry, *1-jettiness DIS event shape: NNLL+NLO results*, Phys. Rev. D **90** (2014), no. 1 014041, [arXiv:1312.0301].

[304] I. W. Stewart, F. J. Tackmann, J. Thaler, C. K. Vermilion, and T. F. Wilkason, *XCone: N-jettiness as an Exclusive Cone Jet Algorithm*, JHEP **11** (2015) 072, [arXiv:1508.01516].

[305] S. Catani, D. de Florian, and M. Grazzini, *Universality of nonleading logarithmic contributions in transverse momentum distributions*, Nucl. Phys. **B596** (2001) 299–312, [hep-ph/0008184].

[306] T. Becher and M. Neubert, *Drell-Yan production at small q_T, transverse parton distributions and the collinear anomaly*, Eur.Phys.J. **C71** (2011) 1665, [arXiv:1007.4005].

[307] M. G. Echevarria, A. Idilbi, and I. Scimemi, *Factorization Theorem For Drell-Yan At Low q_T And Transverse Momentum Distributions On-The-Light-Cone*, JHEP **1207** (2012) 002, [arXiv:1111.4996].

[308] J.-Y. Chiu, A. Jain, D. Neill, and I. Z. Rothstein, *A Formalism for the Systematic Treatment of Rapidity Logarithms in Quantum Field Theory*, JHEP **05** (2012) 084, [arXiv:1202.0814].

[309] S. Catani, L. Cieri, D. de Florian, G. Ferrera, and M. Grazzini, *Universality of transverse-momentum resummation and hard factors at the NNLO*, Nucl. Phys. **B881** (2014) 414–443, [arXiv:1311.1654].

[310] Y. Delenda, R. Appleby, M. Dasgupta, and A. Banfi, *On QCD resummation with k(t) clustering*, JHEP **0612** (2006) 044, [hep-ph/0610242].

[311] C. W. Bauer, A. Hornig, and F. J. Tackmann, *Factorization for generic jet production*, Phys. Rev. D **79** (2009) 114013, [arXiv:0808.2191].

488

[312] R. Kelley, J. R. Walsh, and S. Zuberi, *Disentangling Clustering Effects in Jet Algorithms*, arXiv:1203.2923.

[313] A. Banfi, G. P. Salam, and G. Zanderighi, *NLL+NNLO predictions for jet-veto efficiencies in Higgs-boson and Drell-Yan production*, JHEP **1206** (2012) 159, [arXiv:1203.5773].

[314] T. Becher and M. Neubert, *Factorization and NNLL Resummation for Higgs Production with a Jet Veto*, JHEP **1207** (2012) 108, [arXiv:1205.3806].

[315] A. Banfi, P. F. Monni, G. P. Salam, and G. Zanderighi, *Higgs and Z-boson production with a jet veto*, Phys. Rev. Lett. **109** (2012) 202001, [arXiv:1206.4998].

[316] T. Becher, M. Neubert, and L. Rothen, *Factorization and $N^3 LL_p$ +NNLO predictions for the Higgs cross section with a jet veto*, JHEP **1310** (2013) 125, [arXiv:1307.0025].

[317] D. Y. Shao, C. S. Li, and H. T. Li, *Resummation Prediction on Higgs and Vector Boson Associated Production with a Jet Veto at the LHC*, JHEP **02** (2014) 117, [arXiv:1309.5015].

[318] Y. Li and X. Liu, *High precision predictions for exclusive VH production at the LHC*, JHEP **06** (2014) 028, [arXiv:1401.2149].

[319] P. Jaiswal and T. Okui, *Explanation of the WW excess at the LHC by jet-veto resummation*, Phys. Rev. D **90** (2014), no. 7 073009, [arXiv:1407.4537].

[320] S. Gangal, M. Stahlhofen, and F. J. Tackmann, *Rapidity-Dependent Jet Vetoes*, Phys. Rev. D **91** (2015), no. 5 054023, [arXiv:1412.4792].

[321] T. Becher, R. Frederix, M. Neubert, and L. Rothen, *Automated NNLL + NLO resummation for jet-veto cross sections*, Eur. Phys. J. C **75** (2015), no. 4 154, [arXiv:1412.8408].

[322] S. D. Ellis, A. Hornig, C. Lee, C. K. Vermilion, and J. R. Walsh, *Consistent Factorization of Jet Observables in Exclusive Multijet Cross-Sections*, Phys. Lett. B **689** (2010) 82–89, [arXiv:0912.0262].

[323] T. Becher and X. Garcia i Tormo, *Factorization and resummation for transverse thrust*, JHEP **06** (2015) 071, [arXiv:1502.04136].

[324] M. Procura and I. W. Stewart, *Quark Fragmentation within an Identified Jet*, Phys. Rev. **D81** (2010) 074009, [arXiv:0911.4980]. [Erratum: Phys. Rev.D83,039902(2011)].

[325] X. Liu, *SCET approach to top quark decay*, Phys. Lett. **B699** (2011) 87–92, [arXiv:1011.3872].

489

[326] A. Jain, M. Procura, and W. J. Waalewijn, *Parton Fragmentation within an Identified Jet at NNLL*, *JHEP* **05** (2011) 035, [arXiv:1101.4953].

[327] M. Procura and W. J. Waalewijn, *Fragmentation in Jets: Cone and Threshold Effects*, *Phys. Rev. D* **85** (2012) 114041, [arXiv:1111.6605].

[328] H.-M. Chang, M. Procura, J. Thaler, and W. J. Waalewijn, *Calculating Track-Based Observables for the LHC*, *Phys. Rev. Lett.* **111** (2013) 102002, [arXiv:1303.6637].

[329] C. W. Bauer and E. Mereghetti, *Heavy Quark Fragmenting Jet Functions*, *JHEP* **1404** (2014) 051, [arXiv:1312.5605].

[330] M. Ritzmann and W. J. Waalewijn, *Fragmentation in Jets at NNLO*, *Phys.Rev.* **D90** (2014) 054029, [arXiv:1407.3272].

[331] L. J. Dixon, *Calculating scattering amplitudes efficiently*, hep-ph/9601359.

[332] L. J. Dixon, *A brief introduction to modern amplitude methods*, arXiv:1310.5353.

[333] S. Keppeler and M. Sjodahl, *Orthogonal multiplet bases in SU(Nc) color space*, *JHEP* **1209** (2012) 124, [arXiv:1207.0609].

[334] M. Sjodahl and J. Thorén, *Decomposing color structure into multiplet bases*, arXiv:1507.03814.

[335] N. Kidonakis, G. Oderda, and G. Sterman, *Evolution of color exchange in QCD hard scattering*, *Nucl. Phys. B* **531** (1998) 365–402, [hep-ph/9803241].

[336] F. Maltoni, K. Paul, T. Stelzer, and S. Willenbrock, *Color flow decomposition of QCD amplitudes*, *Phys. Rev. D* **67** (2003) 014026, [hep-ph/0209271].

[337] S. Catani, M. H. Seymour, and Z. Trocsanyi, *Regularization scheme independence and unitarity in QCD cross sections*, *Phys. Rev. D* **55** (1997) 6819–6829, [hep-ph/9610553].

[338] G. 't Hooft and M. Veltman, *Regularization and Renormalization of Gauge Fields*, *Nucl. Phys. B* **44** (1972) 189–213.

[339] Z. Bern and D. A. Kosower, *The Computation of loop amplitudes in gauge theories*, *Nucl. Phys. B* **379** (1992) 451–561.

[340] Z. Bern, A. De Freitas, L. J. Dixon, and H. L. Wong, *Supersymmetric regularization, two-loop QCD amplitudes and coupling shifts*, *Phys. Rev. D* **66** (2002) 085002, [hep-ph/0202271].

[341] W. Siegel, *Supersymmetric Dimensional Regularization via Dimensional Reduction*, *Phys. Lett. B* **84** (1979) 193.

490

[342] A. Broggio, C. Gnendiger, A. Signer, D. Stöckinger, and A. Visconti, *SCET approach to regularization-scheme dependence of QCD amplitudes*, arXiv:1506.05301.

[343] W. B. Kilgore, *Regularization Schemes and Higher Order Corrections*, Phys. Rev. D **83** (2011) 114005, [arXiv:1102.5353].

[344] R. Boughezal, K. Melnikov, and F. Petriello, *The four-dimensional helicity scheme and dimensional reconstruction*, arXiv:1106.5520.

[345] G. Altarelli, G. Curci, G. Martinelli, and S. Petrarca, *QCD Nonleading Corrections to Weak Decays as an Application of Regularization by Dimensional Reduction*, Nucl. Phys. B **187** (1981) 461.

[346] S. Dawson, *Radiative corrections to Higgs boson production*, Nucl. Phys. B **359** (1991) 283–300.

[347] A. Djouadi, M. Spira, and P. M. Zerwas, *Production of Higgs bosons in proton colliders: QCD corrections*, Phys. Lett. B **264** (1991) 440–446.

[348] M. Spira, A. Djouadi, D. Graudenz, and P. M. Zerwas, *Higgs boson production at the LHC*, Nucl. Phys. B **453** (1995) 17–82, [hep-ph/9504378].

[349] R. Harlander and P. Kant, *Higgs production and decay: Analytic results at next-to-leading order QCD*, JHEP **12** (2005) 015, [hep-ph/0509189].

[350] C. Anastasiou, S. Beerli, S. Bucherer, A. Daleo, and Z. Kunszt, *Two-loop amplitudes and master integrals for the production of a Higgs boson via a massive quark and a scalar-quark loop*, JHEP **01** (2007) 082, [hep-ph/0611236].

[351] R. V. Harlander, *Virtual corrections to g g -> H to two loops in the heavy top limit*, Phys. Lett. B **492** (2000) 74–80, [hep-ph/0007289].

[352] R. V. Harlander and K. J. Ozeren, *Top mass effects in Higgs production at next-to-next-to-leading order QCD: virtual corrections*, Phys. Lett. B **679** (2009) 467–472, [arXiv:0907.2997].

[353] A. Pak, M. Rogal, and M. Steinhauser, *Virtual three-loop corrections to Higgs boson production in gluon fusion for finite top quark mass*, Phys. Lett. B **679** (2009) 473–477, [arXiv:0907.2998].

[354] T. Gehrmann, M. Jaquier, E. W. N. Glover, and A. Koukoutsakis, *Two-Loop QCD Corrections to the Helicity Amplitudes for H → 3 partons*, JHEP **02** (2012) 056, [arXiv:1112.3554].

[355] D. Neill, *Analytic Virtual Corrections for Higgs Transverse Momentum Spectrum at $O(\alpha_s^2/m_t^3)$ via Unitarity Methods*, arXiv:0911.2707.

491

[356] S. Dawson and R. Kauffman, *Higgs boson plus multi - jet rates at the SSC*, *Phys.Rev.Lett.* **68** (1992) 2273–2276.

[357] V. Del Duca, W. Kilgore, C. Oleari, C. Schmidt, and D. Zeppenfeld, *Gluon-fusion contributions to H + 2 jet production*, *Nucl. Phys.* B **616** (2001) 367–399, [hep-ph/0108030].

[358] C. F. Berger, V. Del Duca, and L. J. Dixon, *Recursive Construction of Higgs-Plus-Multiparton Loop Amplitudes: The Last of the Phi-nite Loop Amplitudes*, *Phys.Rev.* D **74** (2006) 094021, [hep-ph/0608180].

[359] S. D. Badger, E. W. N. Glover, and K. Risager, *One-loop phi-MHV amplitudes using the unitarity bootstrap*, *JHEP* **07** (2007) 066, [arXiv:0704.3914].

[360] E. W. N. Glover, P. Mastrolia, and C. Williams, *One-loop phi-MHV amplitudes using the unitarity bootstrap: The General helicity case*, *JHEP* **08** (2008) 017, [arXiv:0804.4149].

[361] S. Badger, J. M. Campbell, R. Ellis, and C. Williams, *Analytic results for the one-loop NMHV Hqqgg amplitude*, *JHEP* **0912** (2009) 035, [arXiv:0910.4481].

[362] J. Korner and P. Sieben, *Use of helicity methods in evaluating loop integrals: A QCD example*, *Nucl. Phys.* B **363** (1991) 65–82.

[363] L. Garland, T. Gehrmann, E. N. Glover, A. Koukoutsakis, and E. Remiddi, *The Two loop QCD matrix element for e + e− → 3 jets*, *Nucl.Phys.* **B627** (2002) 107–188, [hep-ph/0112081].

[364] L. Garland, T. Gehrmann, E. N. Glover, A. Koukoutsakis, and E. Remiddi, *Two loop QCD helicity amplitudes for e + e → three jets*, *Nucl.Phys.* **B642** (2002) 227–262, [hep-ph/0206067].

[365] J. van der Bij and E. Glover, *Z BOSON PRODUCTION AND DECAY VIA GLUONS*, *Nucl. Phys.* B **313** (1989) 237.

[366] T. Gehrmann, L. Tancredi, and E. Weihs, *Two-loop QCD helicity amplitudes for g g → Z g and g g → Z γ*, *JHEP* **1304** (2013) 101, [arXiv:1302.2630].

[367] Z. Bern, L. J. Dixon, D. A. Kosower, and S. Weinzierl, *One-loop amplitudes for e+ e- -> anti-q q anti-Q Q*, *Nucl. Phys.* B **489** (1997) 3–23, [hep-ph/9610370].

[368] Z. Bern, A. De Freitas, and L. J. Dixon, *Two loop helicity amplitudes for quark gluon scattering in QCD and gluino gluon scattering in supersymmetric Yang-Mills theory*, *JHEP* **0306** (2003) 028, [hep-ph/0304168].

[369] E. W. N. Glover and M. E. Tejeda-Yeomans, *Two-loop QCD helicity amplitudes for massless quark- massless gauge boson scattering*, *JHEP* **06** (2003) 033, [hep-ph/0304169].

[370] C. Anastasiou, E. N. Glover, C. Oleari, and M. Tejeda-Yeomans, *Two-loop QCD corrections to the scattering of massless distinct quarks, Nucl. Phys.* **B601** (2001) 318–340, [hep-ph/0010212].

[371] C. Anastasiou, E. N. Glover, C. Oleari, and M. Tejeda-Yeomans, *Two loop QCD corrections to massless identical quark scattering, Nucl. Phys.* **B601** (2001) 341–360, [hep-ph/0011094].

[372] E. W. N. Glover, *Two-loop QCD helicity amplitudes for massless quark-quark scattering, JHEP* **04** (2004) 021, [hep-ph/0401119].

[373] A. De Freitas and Z. Bern, *Two-loop helicity amplitudes for quark-quark scattering in QCD and gluino-gluino scattering in supersymmetric Yang-Mills theory, JHEP* **09** (2004) 039, [hep-ph/0409007].

[374] E. N. Glover, C. Oleari, and M. Tejeda-Yeomans, *Two loop QCD corrections to gluon-gluon scattering, Nucl. Phys.* **B605** (2001) 467–485, [hep-ph/0102201].

[375] Z. Bern, A. De Freitas, and L. J. Dixon, *Two loop helicity amplitudes for gluon-gluon scattering in QCD and supersymmetric Yang-Mills theory, JHEP* **0203** (2002) 018, [hep-ph/0201161].

[376] S. Mert Aybat, L. J. Dixon, and G. Sterman, *The two-loop soft anomalous dimension matrix and resummation at next-to-next-to leading pole, Phys. Rev. D* **74** (2006) 074004, [hep-ph/0607309].

[377] S. M. Aybat, L. J. Dixon, and G. F. Sterman, *The Two-loop anomalous dimension matrix for soft gluon exchange, Phys. Rev. Lett.* **97** (2006) 072001, [hep-ph/0606254].

[378] Ø. Almelid, C. Duhr, and E. Gardi, *Three-loop corrections to the soft anomalous dimension in multi-leg scattering,* arXiv:1507.00047.

[379] E. Gardi and L. Magnea, *Infrared singularities in QCD amplitudes, Nuovo Cim.* **C32N5-6** (2009) 137–157, [arXiv:0908.3273].

[380] E. Gardi and L. Magnea, *Factorization constraints for soft anomalous dimensions in QCD scattering amplitudes, JHEP* **03** (2009) 079, [arXiv:0901.1091].

[381] L. J. Dixon, E. Gardi, and L. Magnea, *On soft singularities at three loops and beyond, JHEP* **02** (2010) 081, [arXiv:0910.3653].

[382] T. Becher and M. Neubert, *Infrared singularities of scattering amplitudes in perturbative QCD, Phys. Rev. Lett.* **102** (2009) 162001, [arXiv:0901.0722].

[383] T. Becher and M. Neubert, *On the Structure of Infrared Singularities of Gauge-Theory Amplitudes, JHEP* **06** (2009) 081, [arXiv:0903.1126].

[384] V. Del Duca, C. Duhr, E. Gardi, L. Magnea, and C. D. White, *The Infrared structure of gauge theory amplitudes in the high-energy limit*, JHEP **1112** (2011) 021, [arXiv:1109.3581].

[385] V. Del Duca, C. Duhr, E. Gardi, L. Magnea, and C. D. White, *An infrared approach to Reggeization*, Phys. Rev. D **85** (2012) 071104, [arXiv:1108.5947].

[386] V. Del Duca, G. Falcioni, L. Magnea, and L. Vernazza, *High-energy QCD amplitudes at two loops and beyond*, Phys. Lett. **B732** (2014) 233–240, [arXiv:1311.0304].

[387] S. Caron-Huot, *When does the gluon reggeize?*, arXiv:1309.6521.

[388] S. Catani and M. Seymour, *The Dipole formalism for the calculation of QCD jet cross-sections at next-to-leading order*, Phys.Lett. **B378** (1996) 287–301, [hep-ph/9602277].

[389] S. Catani and M. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, Nucl.Phys. **B485** (1997) 291–419, [hep-ph/9605323].

[390] E. Gerwick, S. Hoeche, S. Marzani, and S. Schumann, *Soft evolution of multi-jet final states*, JHEP **1502** (2015) 106, [arXiv:1411.7325].

[391] S. Plätzer, *Summing Large-N Towers in Colour Flow Evolution*, Eur. Phys. J. **C74** (2014), no. 6 2907, [arXiv:1312.2448].

[392] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, *The Quark Beam Function at NNLL*, JHEP **1009** (2010) 005, [arXiv:1002.2213].

[393] **ATLAS Collaboration** Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys.Lett. **B716** (2012) 1–29, [arXiv:1207.7214].

[394] **CMS Collaboration** Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys.Lett. **B716** (2012) 30–61, [arXiv:1207.7235].

[395] M. Duhrssen, S. Heinemeyer, H. Logan, D. Rainwater, G. Weiglein, et al., *Extracting Higgs boson couplings from CERN LHC data*, Phys.Rev. **D70** (2004) 113009, [hep-ph/0406323].

[396] R. Lafaye, T. Plehn, M. Rauch, D. Zerwas, and M. Duhrssen, *Measuring the Higgs Sector*, JHEP **0908** (2009) 009, [arXiv:0904.3866].

[397] F. Bonnet, M. Gavela, T. Ota, and W. Winter, *Anomalous Higgs couplings at the LHC, and their theoretical interpretation*, Phys.Rev. **D85** (2012) 035016, [arXiv:1105.5140].

494

[398] **LHC Higgs Cross Section Working Group** Collaboration, S. Heinemeyer et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, arXiv:1307.1347.

[399] D. Carmi, A. Falkowski, E. Kuflik, T. Volansky, and J. Zupan, *Higgs After the Discovery: A Status Report*, JHEP **1210** (2012) 196, [arXiv:1207.1718].

[400] D. Carmi, A. Falkowski, E. Kuflik, and T. Volansky, *Interpreting LHC Higgs Results from Natural New Physics Perspective*, JHEP **1207** (2012) 136, [arXiv:1202.3144].

[401] A. Djouadi, *Precision Higgs coupling measurements at the LHC through ratios of production cross sections*, Eur.Phys.J. **C73** (2013) 2498, [arXiv:1208.3436].

[402] C. Englert, M. Spannowsky, and C. Wymant, *Partially (in)visible Higgs decays at the LHC*, Phys.Lett. **B718** (2012) 538–544, [arXiv:1209.0494].

[403] M. Klute, R. Lafaye, T. Plehn, M. Rauch, and D. Zerwas, *Measuring Higgs Couplings from LHC Data*, Phys.Rev.Lett. **109** (2012) 101801, [arXiv:1205.2699].

[404] B. A. Dobrescu and J. D. Lykken, *Coupling spans of the Higgs-like boson*, JHEP **1302** (2013) 073, [arXiv:1210.3342].

[405] T. Plehn and M. Rauch, *Higgs Couplings after the Discovery*, Europhys.Lett. **100** (2012) 11002, [arXiv:1207.6108].

[406] T. Corbett, O. Eboli, J. Gonzalez-Fraile, and M. Gonzalez-Garcia, *Robust Determination of the Higgs Couplings: Power to the Data*, Phys.Rev. **D87** (2013) 015022, [arXiv:1211.4580].

[407] G. Belanger, B. Dumont, U. Ellwanger, J. Gunion, and S. Kraml, *Higgs Couplings at the End of 2012*, JHEP **1302** (2013) 053, [arXiv:1212.5244].

[408] M. Farina, C. Grojean, F. Maltoni, E. Salvioni, and A. Thamm, *Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson*, JHEP **1305** (2013) 022, [arXiv:1211.3736].

[409] B. Batell, S. Gori, and L.-T. Wang, *Higgs Couplings and Precision Electroweak Data*, JHEP **1301** (2013) 139, [arXiv:1209.6382].

[410] J. Espinosa, C. Grojean, M. Muhlleitner, and M. Trott, *First Glimpses at Higgs' face*, JHEP **1212** (2012) 045, [arXiv:1207.1717].

[411] J. R. Espinosa, M. Muhlleitner, C. Grojean, and M. Trott, *Probing for Invisible Higgs Decays with Global Fits*, JHEP **1209** (2012) 126, [arXiv:1205.6790].

[412] S. Banerjee, S. Mukhopadhyay, and B. Mukhopadhyaya, *New Higgs interactions and recent data from the LHC and the Tevatron*, JHEP **1210** (2012) 062, [arXiv:1207.3588].

[413] V. Barger, M. Ishida, and W.-Y. Keung, *Total Width of 125 GeV Higgs Boson*, *Phys.Rev.Lett.* **108** (2012) 261801, [arXiv:1203.3456].

[414] P. P. Giardino, K. Kannike, M. Raidal, and A. Strumia, *Is the resonance at 125 GeV the Higgs boson?*, *Phys.Lett.* **B718** (2012) 469–474, [arXiv:1207.1347].

[415] P. P. Giardino, K. Kannike, M. Raidal, and A. Strumia, *Reconstructing Higgs boson properties from the LHC and Tevatron data*, *JHEP* **1206** (2012) 117, [arXiv:1203.4254].

[416] P. P. Giardino, K. Kannike, I. Masina, M. Raidal, and A. Strumia, *The universal Higgs fit*, arXiv:1303.3570.

[417] **ATLAS Collaboration** Collaboration, G. Aad et al., *Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC*, *Phys.Lett.* **B726** (2013) 88–119, [arXiv:1307.1427].

[418] J. S. Gainer, J. Lykken, K. T. Matchev, S. Mrenna, and M. Park, *Geolocating the Higgs Boson Candidate at the LHC*, *Phys.Rev.Lett.* **111** (2013) 041801, [arXiv:1304.4936].

[419] G. Belanger, B. Dumont, U. Ellwanger, J. Gunion, and S. Kraml, *Global fit to Higgs signal strengths and couplings and implications for extended Higgs sectors*, *Phys.Rev.* **D88** (2013) 075008, [arXiv:1306.2941].

[420] G. Belanger, B. Dumont, U. Ellwanger, J. Gunion, and S. Kraml, *Status of invisible Higgs decays*, *Phys.Lett.* **B723** (2013) 340–347, [arXiv:1302.5694].

[421] J. Ellis and T. You, *Updated Global Analysis of Higgs Couplings*, *JHEP* **1306** (2013) 103, [arXiv:1303.3879].

[422] K. Cranmer, S. Kreiss, D. Lopez-Val, and T. Plehn, *A Novel Approach to Higgs Coupling Measurements*, arXiv:1401.0080.

[423] Y. Chen, E. Di Marco, J. Lykken, M. Spiropulu, R. Vega-Morales, et al., *8D Likelihood Effective Higgs Couplings Extraction Framework in the Golden Channel*, arXiv:1401.2077.

[424] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model*, *Phys.Rept.* **459** (2008) 1–241, [hep-ph/0503173].

[425] C. Burgess, M. Pospelov, and T. ter Veldhuis, *The Minimal model of nonbaryonic dark matter: A Singlet scalar*, *Nucl.Phys.* **B619** (2001) 709–728, [hep-ph/0011335].

[426] B. Patt and F. Wilczek, *Higgs-field portal into hidden sectors*, hep-ph/0605188.

[427] X.-G. He and J. Tandean, *Hidden Higgs Boson at the LHC and Light Dark Matter Searches*, *Phys.Rev.* **D84** (2011) 075018, [arXiv:1109.1277].

[428] M. Raidal and A. Strumia, *Hints for a non-standard Higgs boson from the LHC*, *Phys.Rev.* **D84** (2011) 077701, [arXiv:1108.4903].

[429] C. Englert, T. Plehn, D. Zerwas, and P. M. Zerwas, *Exploring the Higgs portal*, *Phys.Lett.* **B703** (2011) 298–305, [arXiv:1106.3097].

[430] R. Barbieri, T. Gregoire, and L. J. Hall, *Mirror world at the large hadron collider*, hep-ph/0509242.

[431] S. Gangal and F. J. Tackmann, *Next-to-leading-order uncertainties in Higgs+2 jets from gluon fusion*, *Phys.Rev.* **D87** (2013), no. 9 093008, [arXiv:1302.5437].

[432] L. J. Dixon and Y. Li, *Bounding the Higgs Boson Width Through Interferometry*, *Phys.Rev.Lett.* **111** (2013) 111802, [arXiv:1305.3854].

[433] J. M. Campbell, R. K. Ellis, and C. Williams, *Bounding the Higgs width at the LHC using full analytic results for gg → 2e2μ*, arXiv:1311.3589.

[434] J. M. Campbell, R. K. Ellis, and C. Williams, *Bounding the Higgs width at the LHC: complementary results from H → WW*, arXiv:1312.1628.

[435] **CMS Collaboration** Collaboration, C. Collaboration, *Search for an Invisible Higgs Boson*, .

[436] **CMS Collaboration** Collaboration, C. Collaboration, *Search for invisible Higgs produced in association with a Z boson*, .

[437] Y. Bai, P. Draper, and J. Shelton, *Measuring the Invisible Higgs Width at the 7 and 8 TeV LHC*, *JHEP* **1207** (2012) 192, [arXiv:1112.4496].

[438] A. Djouadi, A. Falkowski, Y. Mambrini, and J. Quevillon, *Direct Detection of Higgs-Portal Dark Matter at the LHC*, *Eur.Phys.J.* **C73** (2013) 2455, [arXiv:1205.3169].

[439] J. M. Campbell, R. K. Ellis, and C. Williams, *Gluon-Gluon Contributions to W+ W- Production and Higgs Interference Effects*, *JHEP* **1110** (2011) 005, [arXiv:1107.5569].

[440] N. Kauer, *Signal-background interference in gg →H→ VV*, *PoS* **RADCOR2011** (2011) 027, [arXiv:1201.1667].

[441] G. Passarino, *Higgs Interference Effects in gg → ZZ and their Uncertainty*, *JHEP* **1208** (2012) 146, [arXiv:1206.3824].

[442] N. Kauer, *Interference effects for H → WW/ZZ → ℓ$\bar{\nu}_\ell$ℓ$\bar{\nu}_\ell$ searches in gluon fusion at the LHC*, *JHEP* **1312** (2013) 082, [arXiv:1310.7011].

[443] C. Uhlemann and N. Kauer, *Narrow-width approximation accuracy*, *Nucl.Phys.* **B814** (2009) 195–211, [arXiv:0807.4112].

[444] N. Kauer, *Inadequacy of zero-width approximation for a light Higgs boson signal*, *Mod.Phys.Lett.* **A28** (2013) 1330015, [arXiv:1305.2092].

[445] C. Englert and M. Spannowsky, *Limitations and Opportunities of Off-Shell Coupling Measurements*, arXiv:1405.0285.

[446] **CMS Collaboration** Collaboration, C. Collaboration, *Constraints on the Higgs boson width from off-shell production and decay to ZZ to llll and llνν*, .

[447] B. Coleppa, T. Mandal, and S. Mitra, *Coupling Extraction From Off-Shell Cross-sections*, arXiv:1401.4039.

[448] J. M. Henn, K. Melnikov, and V. A. Smirnov, *Two-loop planar master integrals for the production of off-shell vector bosons in hadron collisions*, arXiv:1402.7078.

[449] F. Caola, J. M. Henn, K. Melnikov, and V. A. Smirnov, *Non-planar master integrals for the production of two off-shell vector bosons in collisions of massless partons*, arXiv:1404.5590.

[450] S. Dittmaier, S. Dittmaier, C. Mariotti, G. Passarino, R. Tanaka, et al., *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, arXiv:1201.3084.

[451] I. W. Stewart and F. J. Tackmann, *Theory Uncertainties for Higgs and Other Searches Using Jet Bins*, *Phys. Rev.* **D 85** (2012) 034011, [arXiv:1107.2117].

[452] S. Alioli and J. R. Walsh, *Jet Veto Clustering Logarithms Beyond Leading Order*, *JHEP* **1403** (2014) 119, [arXiv:1311.5234].

[453] J.-y. Chiu, A. Jain, D. Neill, and I. Z. Rothstein, *The Rapidity Renormalization Group*, *Phys.Rev.Lett.* **108** (2012) 151601, [arXiv:1104.0881].

[454] S. Fleming, A. K. Leibovich, and T. Mehen, *Resummation of Large Endpoint Corrections to Color-Octet J/ψ Photoproduction*, *Phys. Rev.* **D 74** (2006) 114004, [hep-ph/0607121].

[455] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, *Eur. Phys. J.* **C 63** (2009) 189–285, [arXiv:0901.0002].

[456] R. Boughezal, X. Liu, F. Petriello, F. J. Tackmann, and J. R. Walsh, *Combining Resummed Higgs Predictions Across Jet Bins*, *Phys.Rev.* **D89** (2014) 074044, [arXiv:1312.4535].

[457] J. Campbell, K. Hatakeyama, J. Huston, F. Petriello, J. R. Andersen, et al., *Report of the Snowmass 2013 energy frontier QCD working group*, arXiv:1310.5189.

[458] J. M. Campbell and R. Ellis, *MCFM for the Tevatron and the LHC*, *Nucl.Phys.Proc.Suppl.* **205-206** (2010) 10–15, [arXiv:1007.3492].

[459] A. Djouadi, J. Kalinowski, and M. Spira, *HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension*, *Comput.Phys.Commun.* **108** (1998) 56–74, [hep-ph/9704448].

[460] T. Hahn, *Automatic loop calculations with FeynArts, FormCalc, and LoopTools*, *Nucl.Phys.Proc.Suppl.* **89** (2000) 231–236, [hep-ph/0005029].

[461] T. Hahn, *CUBA: A Library for multidimensional numerical integration*, *Comput. Phys. Commun.* **168** (2005) 78–95, [hep-ph/0404043].

[462] A. Banfi, P. F. Monni, and G. Zanderighi, *Quark masses in Higgs production with a jet veto*, *JHEP* **1401** (2014) 097, [arXiv:1308.4634].

[463] M. Bonvini, F. Caola, S. Forte, K. Melnikov, and G. Ridolfi, *Signal-background interference effects for $gg \to H \to W^+W^-$ beyond leading order*, *Phys.Rev.* **D88** (2013) 034032, [arXiv:1304.3053].

[464] T. A. collaboration, *Search for a high-mass Higgs boson in the $H \to WW \to l\nu l\nu$ decay channel with the ATLAS detector using 21 fb^{-1} of proton-proton collision data,*

[465] **CMS Collaboration** Collaboration, S. Chatrchyan et al., *Search for a standard-model-like Higgs boson with a mass in the range 145 to 1000 GeV at the LHC*, *Eur.Phys.J.* **C73** (2013) 2469, [arXiv:1304.0213].

[466] **ATLAS Collaboration** Collaboration, *Measurements of the properties of the Higgs-like boson in the $WW^{(*)} \to l\nu l\nu$ decay channel with the ATLAS detector using 25 fb^{-1} of proton-proton collision data, .*

[467] **CMS Collaboration** Collaboration, *Evidence for a particle decaying to W+W- in the fully leptonic final state in a standard model Higgs boson search in pp collisions at the LHC, .*

[468] I. W. Stewart and C. W. Bauer, "Lectures on the soft-collinear effective theory."

[469] T. Becher, A. Broggio, and A. Ferroglia, *Introduction to Soft-Collinear Effective Theory*, arXiv:1410.1892.

[470] N. A. Sveshnikov and F. V. Tkachov, *Jets and quantum field theory*, *Phys. Lett. B* **382** (1996) 403–408, [hep-ph/9512370].

[471] G. P. Korchemsky, G. Oderda, and G. Sterman, *Power corrections and nonlocal operators*, hep-ph/9708346.

499

[472] C. Lee and G. F. Sterman, *Momentum Flow Correlations from Event Shapes: Factorized Soft Gluons and Soft-Collinear Effective Theory*, Phys. Rev. D **75** (2007) 014022, [hep-ph/0611061].

[473] C. W. Bauer, S. P. Fleming, C. Lee, and G. Sterman, *Factorization of e^+e^- Event Shape Distributions with Hadronic Final States in Soft Collinear Effective Theory*, Phys. Rev. D **78** (2008) 034027, [arXiv:0801.4569].

[474] C. W. Bauer, C. Lee, A. V. Manohar, and M. B. Wise, *Enhanced nonperturbative effects in Z decays to hadrons*, Phys. Rev. D **70** (2004) 034014, [hep-ph/0309278].

[475] A. V. Manohar, *Deep inelastic scattering as $x \to 1$ using soft-collinear effective theory*, Phys. Rev. D **68** (2003) 114019, [hep-ph/0309176].

[476] D. A. Kosower and P. Uwer, *One loop splitting amplitudes in gauge theory*, Nucl.Phys. **B563** (1999) 477–505, [hep-ph/9903515].

[477] R. K. Ellis, D. A. Ross, and A. E. Terrano, *The Perturbative Calculation of Jet Structure in e+ e- Annihilation*, Nucl. Phys. **B178** (1981) 421.

[478] Z. Bern, L. J. Dixon, and D. A. Kosower, *Two-loop $g \longrightarrow gg$ splitting amplitudes in QCD*, JHEP **08** (2004) 012, [hep-ph/0404293].

[479] S. Badger and E. N. Glover, *Two loop splitting functions in QCD*, JHEP **0407** (2004) 040, [hep-ph/0405236].

[480] T. Huber and D. Maitre, *HypExp: A Mathematica package for expanding hypergeometric functions around integer-valued parameters*, Comput.Phys.Commun. **175** (2006) 122–144, [hep-ph/0507094].

[481] T. Huber and D. Maitre, *HypExp 2, Expanding Hypergeometric Functions about Half-Integer Parameters*, Comput.Phys.Commun. **178** (2008) 755–776, [arXiv:0708.2443].

[482] F. A. Berends and W. Giele, *Multiple Soft Gluon Radiation in Parton Processes*, Nucl.Phys. **B313** (1989) 595.

[483] S. Catani and M. Grazzini, *The soft gluon current at one loop order*, Nucl.Phys. **B591** (2000) 435–454, [hep-ph/0007142].

[484] C. Duhr and T. Gehrmann, *The two-loop soft current in dimensional regularization*, Phys.Lett. **B727** (2013) 452–455, [arXiv:1309.4393].

[485] Y. Li and H. X. Zhu, *Single soft gluon emission at two loops*, JHEP **1311** (2013) 080, [arXiv:1309.4391].

[486] G. Altarelli, R. K. Ellis, and G. Martinelli, *Large Perturbative Corrections to the Drell-Yan Process in QCD*, Nucl. Phys. B **157** (1979) 461.

[487] T. Matsuura, S. C. van der Marck, and W. L. van Neerven, *The Calculation of the Second Order Soft and Virtual Contributions to the Drell-Yan Cross-Section*, Nucl. Phys. B **319** (1989) 570.

[488] L. G. Almeida, S. D. Ellis, C. Lee, G. Sterman, I. Sung, et al., *Comparing and counting logs in direct and effective methods of QCD resummation*, JHEP **1404** (2014) 174, [arXiv:1401.4460].

[489] A. Banfi, H. McAslan, P. F. Monni, and G. Zanderighi, *A general method for the resummation of event-shape distributions in e^+e^- annihilation*, JHEP **05** (2015) 102, [arXiv:1412.2126].

[490] Y. Sakaki, *Evolution variable dependence of jet substructure*, arXiv:1506.04811.

[491] H. Elvang and Y.-t. Huang, *Scattering Amplitudes*, arXiv:1308.1697.

[492] S. Catani, *The Singular behavior of QCD amplitudes at two loop order*, Phys. Lett. B**427** (1998) 161–171, [hep-ph/9802439].

[493] S. J. Parke and T. R. Taylor, *An Amplitude for n Gluon Scattering*, Phys. Rev. Lett. **56** (1986) 2459.

[494] F. Retiere et al., *Characterization of Multi Pixel Photon Counters for T2K Near Detector*, Nucl. Instrum. Meth. **A610** (2009) 378–380.

[495] C. S. Chiu, C. Ignarra, L. Bugel, H. Chen, J. M. Conrad, B. J. P. Jones, T. Katori, and I. Moult, *Environmental Effects on TPB Wavelength-Shifting Coatings*, JINST **7** (2012) P07007, [arXiv:1204.5762].

[496] O. V. Tarasov, A. A. Vladimirov, and A. Y. Zharkov, *The Gell-Mann-Low Function of QCD in the Three Loop Approximation*, Phys. Lett. B **93** (1980) 429–432.

[497] S. A. Larin and J. A. M. Vermaseren, *The three-loop QCD β function and anomalous dimensions*, Phys. Lett. B **303** (1993) 334–336, [hep-ph/9302208].

[498] G. P. Korchemsky and A. V. Radyushkin, *Renormalization of the Wilson Loops Beyond the Leading Order*, Nucl. Phys. B **283** (1987) 342–364.

[499] S. Moch, J. A. M. Vermaseren, and A. Vogt, *The three-loop splitting functions in QCD: The non-singlet case*, Nucl. Phys. B **688** (2004) 101–134, [hep-ph/0403192].