Exploiting jet binning to identify the initial state of high-mass resonances

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Exploiting jet binning to identify the initial state of high-mass resonances

Markus A. Ebert,1 Stefan Liebler,1 Ian Moult,2 Iain W. Stewart,2 Frank J. Tackmann,1 Kerstin Tackmann,1 and Lisa Zeune3

1Deutsches Elektronen-Synchrotron (DESY), D-22607 Hamburg, Germany
2Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
3Nikhef, Theory Group, Science Park 105, 1098 XG Amsterdam, The Netherlands

(Received 7 June 2016; published 28 September 2016)

If a new high-mass resonance is discovered at the Large Hadron Collider, model-independent techniques to identify the production mechanism will be crucial to understand its nature and effective couplings to Standard Model particles. We present a powerful and model-independent method to infer the initial state in the production of any high-mass color-singlet system by using a tight veto on accompanying hadronic jets to divide the data into two mutually exclusive event samples (jet bins). For a resonance of several hundred GeV, the jet binning cut needed to discriminate quark and gluon initial states is in the experimentally accessible range of several tens of GeV. It also yields comparable cross sections for both bins, making this method viable already with the small event samples available shortly after a discovery. Theoretically, the method is made feasible by utilizing an effective field theory setup to compute the jet cut dependence precisely and model independently and to systematically control all sources of theoretical uncertainties in the jet binning, as well as their correlations. We use a 750 GeV scalar resonance as an example to demonstrate the viability of our method.

DOI: 10.1103/PhysRevD.94.051901

1. INTRODUCTION

The increased center-of-mass energy of the Large Hadron Collider (LHC) significantly enhances the sensitivity for the discovery of new heavy particles. Should a new high-mass state be found, a key goal will be to identify its production mechanism.

It is well known that the different patterns of initial-state radiation (ISR) for gluon- and quark-induced processes provide in principle a way to discriminate between these initial states. Typically, methods to exploit this fact require a substantial amount of data for the precise measurement of shapes of differential distributions. In this paper, we show that for any high-mass color-singlet system, the measurement of just two cross sections, namely dividing the data into events with and without additional hadronic jets in the final state, provides a strong discrimination between production mechanisms, which is furthermore experimentally accessible with event samples of limited size. The method is also theoretically clean, as it is both model independent and has well-controlled theory uncertainties.

As a concrete example, we investigate a color-singlet resonance with a mass of 750 GeV. The ATLAS and CMS experiments have recently reported some deviation from the background expectation in the diphoton invariant mass spectrum around 750 GeV [1,2]. Assuming the deviation to be a first sign of a new particle, a large number of proposals on its interpretation and possible property studies have been made [3]. Exploratory studies of the initial state have utilized the luminosity ratio between 8 and 13 TeV (which is limited by the available 8 TeV data), the transverse momentum and rapidity distribution of the new state [4], multiplicity and kinematic distributions of hadronic jets [5], and b-tagging [4,6]. Different techniques for tagging the initial state have been studied earlier, see e.g. Refs. [7–12]. Beyond its viability for small data sets, our method offers several additional advantages. Compared to considering additional jets at high \( p_T^{\text{jet}} \) [5,6,13], the low \( p_T^{\text{jet}} \)-range we exploit has more discrimination power and is more model independent. Compared to the diphoton \( p_T \) spectrum, the \( p_T^{\text{jet}} \) of hadronic jets provides a more direct measure of ISR, making it insensitive to the possibility of more complicated decays of the resonance, for example, three-body [14,15] or cascade decays [16–21]. Our method is also unaffected by limited experimental acceptance for photons, which for example hinders fully exploiting the diphoton rapidity distribution to discriminate valence quarks by their different parton distribution function (PDF) shapes.

The 0-jet cross section is defined by requiring that all accompanying jets have \( p_T^{\text{jet}} \lesssim p_T^{\text{cut}} \). The QCD dynamics of low-\( p_T \) radiation produced in association with a hard scattering process into a final state \( F \) with total invariant mass \( m(F) = m_X \) can be described using the soft collinear effective theory (SCET) [22–25]. At the scale \( \mu \sim p_T^{\text{cut}} \ll m_X \), the leading effective field theory (EFT) Lagrangian has the form (see e.g. [26–28])

\[
\mathcal{L}_{\text{eff}}(p_T^{\text{cut}}) = \mathcal{L}_{\text{SCET}} + c_{gh}^{(1)} B_{\gamma}^{(1)} B_{\gamma}^{(1)} F + \sum_{q} c_{qF}^{(1)} X_{qF}^{(1)} X_{qF}^{(1)} F.
\]
Here, $\mathcal{L}_{\text{SCET}}$ is the universal Lagrangian encoding the interactions of soft and collinear quarks and gluons. The gauge-invariant operators $B_{\mu}X_{\nu}$ and $X_{\lambda}q_{\lambda}$ describe the annihilation of energetic gluons or quarks $q = u, d, s, c, b$ along the beam directions, $n = (1, 0, 0)$ and $n = (1, -1, 0)$, with helicities $\lambda_1$ and $\lambda_2$ (implicitly summed over), and $X$ collects all fields required to produce $F$. All hard degrees of freedom are integrated out, including quarks and gluons of virtuality $\sim m_X$ as well as any intermediate new heavy degrees of freedom leading to $F$.

We stress that $\mathcal{L}_{\text{eff}}(m_X)$ provides a completely model-independent description of the small-$p_T^\text{cut}$ region, up to power corrections suppressed by $(p_T^\text{cut}/m_X)^2$. It is valid for any produced color-singlet system $X$ leading to $F$, for example the decay of a finite-width resonance, the Standard Model background $pp \to F$, and even the signal-background interference [29], as all of the dynamics of the production and decay of $X$ are contained in the hard Wilson coefficients $c_{gqF}^{1/2}$ and $c_{gqF}^{1/2}$ and because the leading perturbative SCET dynamics are insensitive to the helicity structure of the operators and the details of $\mathcal{F}$. The 0-jet cross section thus only depends on the hard coefficients

$$
|c_{ij}|^2 = \int d\Phi_F \sum_{\lambda_1\lambda_2} |c_{ij}^{\lambda_1\lambda_2}(\Phi_F)|^2,
$$

where the integral is over the final-state phase space for $F$ including any kinematic selection cuts. As a result, the $p_T^\text{cut}$ dependence is independent of any details of $F$ and in particular also the spin of $X$. (See Ref. [30] for a detailed analysis in a specific new-physics context.)

To predict the inclusive cross section $pp \to X \to F$ we need to know the Lagrangian at the scale $\mu \sim m_X$, which contains the full QCD Lagrangian plus the (effective) interactions of $X$ with quarks and gluons. This becomes somewhat more model dependent, and requires, for example, specifying the spin of $X$. For our concrete study we take $X$ to be a scalar, coupling to gluons and quarks via the effective Wilson coefficients $C_g$ and $C_q$ as

$$
\mathcal{L}_{\text{eff}}(m_X) \supset -\frac{C_g}{1 \text{ TeV}} \alpha_s G^{\mu\nu} G_{\mu\nu} X - \sum_q C_q q q X, \quad \text{(3)}
$$

where $G^{\mu\nu}$ is the gluon field strength and $\alpha_s$ is the strong coupling. (We assume that $X$ does not couple directly to top quarks, as this would have shown up in $t\bar{t}$ production.) Comparing quark and gluon luminosities as is often done is equivalent to using Eq. (3) at leading order (LO). With Eq. (3) specified, we can now match it onto Eq. (1) and compute $|c_{qgF}|^2$ and $|c_{qgF}|^2$.

For our purposes, any model can be represented by Eq. (3) at leading order in $\alpha_s(m_X)$. Treated as an EFT, Eq. (3) is a priori only correct to $\mathcal{O}(m_X/\Lambda)$, where $\Lambda$ is the mass scale of additional heavy degrees of freedom that induce the effective interactions of $X$. For example, the possibility of real QCD radiation from internal heavy states is not captured by Eq. (3). However, even for $\Lambda \sim m_X$ hard emissions only affect the inclusive cross section by $\mathcal{O}(\alpha_s(m_X))$, while emissions below the scale $p_T^\text{cut}$ are power suppressed. Similarly, a different choice of $\mathcal{L}_{\text{eff}}(m_X)$ in Eq. (3) (e.g. for a spin-2 resonance) changes the inclusive cross section and the matching in Eq. (4) only by terms of $\mathcal{O}(\alpha_s(m_X))$, i.e., at the 10%-20% level. The crucial point is that the $p_T^\text{cut}$ dependence for $p_T^\text{cut} \ll m_X$ is described by Eq. (1). Hence, for more complicated scenarios than the one considered here, our main conclusions regarding the initial-state discrimination are unaffected as they rest on the dynamics at the scale $\mu \sim p_T^\text{cut}$, which is described model independently.

II. CALCULATIONAL SETUP

Considering for simplicity the narrow-width approximation, we have

$$
|c_{qgF}(\mu_H)|^2 = B(X \to F)|C_q(\mu_H)/(1 + \cdots)|^2,
$$

$$
|c_{gqF}(\mu_H)|^2 = B(X \to F)|\alpha_s(\mu_H)C_g(\mu_H)/(1 + \cdots)|^2, \quad \text{(4)}
$$

where $\mu_H \sim m_X$ is the hard matching scale, and the ellipses indicate the $\alpha_s(\mu_H)$ corrections from hard virtual QCD emissions. The branching ratio $B \equiv B(X \to F)$ also depends on all Wilson coefficients $C_i$, but will drop out in our final analysis.

The jet cross sections we consider are given by

$$
\sigma_{00} = |C_q|^2 \sigma_{00}^q + \sum_q |C_q|^2 \sigma_{00}^q,
$$

$$
\sigma_0(p_T^\text{cut}) = |C_g|^2 \sigma_0(p_T^\text{cut}) + \sum_q |C_q|^2 \sigma_0^q(p_T^\text{cut}),
$$

$$
\sigma_{21}(p_T^\text{cut}) = |C_g|^2 \sigma_{21}(p_T^\text{cut}) + \sum_q |C_q|^2 \sigma_{21}^q(p_T^\text{cut}), \quad \text{(5)}
$$

where $\sigma_{00} = \sigma_0(p_T^\text{cut}) + \sigma_{21}(p_T^\text{cut})$. We take $C_i \equiv C_i(m_X)$ as the unknown parameters to be determined from the data. Their evolution from the fixed input scale $m_X$ to the hard matching scale $\mu_H$ is included in the $\sigma_m$ in Eq. (5), so they are defined to be scale independent to all orders. The $\sigma_{00}$ and $\sigma_{20}$ are the inclusive cross sections that follow from the $ggX$ and $q\bar{q}X$ operators in $\mathcal{L}_{\text{eff}}(m_X)$ in Eq. (3). The inclusive 1-jet cross sections are computed as $\sigma_{12}(p_T^\text{cut}) = \sigma_{12}^q - \sigma_{01}^q(p_T^\text{cut})$. The 0-jet cross sections contain large Sudakov logarithms of $p_T^\text{cut}/m_X$, which are resummed utilizing the $p_T^\text{cut}$ resummation framework of Refs. [31,32] based on SCET (see also Refs. [33–36]). They are given by
EXPLOITING JET BINNING TO IDENTIFY THE …

\[ |C_\ell|^2 \sigma_0^g(p_{T}^{cut}) = \frac{\pi}{4E_{\text{cm}}^2} \frac{m_X^2}{2 \text{TeV}^2} \int dYB_{\ell}(p_{T}^{cut}, \mu) \]
\[ \times B_{\ell}(p_{T}^{cut}, \mu) S_{gg}(p_{T}^{cut}, \mu) + \sigma_0^{\text{nons}}(p_{T}^{cut}), \]
\[ |C_q|^2 \sigma_0^q(p_{T}^{cut}) = \frac{\pi}{6E_{\text{cm}}^2} \int dYB_q(p_{T}^{cut}, \mu) \]
\[ \times B_q(p_{T}^{cut}, \mu) S_{qq}(p_{T}^{cut}, \mu) + \sigma_0^{\text{nons}}(p_{T}^{cut}). \]

(6)

The \( B_i \) are quark and gluon beam functions, which describe the dynamics of collinear radiation along the beam directions. In general, the type of the incoming parton is changed by both collinear PDF evolution and fixed-order corrections, so beyond LO an operator in the Lagrangian receives contributions from all PDFs. However, in the 0-jet cross section, both of these effects only occur up to the scale \( p_{T}^{cut} \) and are contained in the beam functions. Above the scale \( p_{T}^{cut} \), the parton type of the initial state is uniquely defined and matches between the beam function and the operator in the Lagrangian [26]. Similarly, the dynamics of wide-angle soft radiation, described by the soft functions \( S_{qq(qq)} \), is unique to the parton type and does not change it. The fact that the jet veto freezes the initial-state parton type at the scale \( p_{T}^{cut} \) is what lends our method its strong discrimination power, as it provides a large energy range between \( p_{T}^{cut} \) and \( m_X \) where the initial state evolves without changing its type.

We calculate \( \sigma_0^i \) to next-to-leading order (NLO) in \( \alpha_s \), and \( \sigma_0^q(p_{T}^{cut}) \) is resummed to NLL’ + NLO order. Due to the substantially larger uncertainties for gluons, we include the full next-to-next-to-leading order (NNLO) corrections for \( \sigma_0^g \), and \( \sigma_0^q(p_{T}^{cut}) \) is resummed to NNLL’ + NNLO [32]. The inclusive cross sections are obtained with SUSHI 1.6.0 [37–41].

The nonsingular corrections \( \sigma_0^{\text{nons}}(p_{T}^{cut}) \) in Eq. (6) contain the power corrections starting at \( (p_{T}^{cut}/m_X)^2 \). They ensure that \( \sigma_0^q(p_{T}^{cut}) \) smoothly matches onto \( \sigma_0^g \) for large \( p_{T}^{cut} \), and are correspondingly included to NLO for quarks and NNLO for gluons. They are extracted from the fixed-order \( p_{T}^{cut} \) spectra predicted by Eq. (3), obtained from SUSHI for \( q\bar{q}X \) and MCFM [42,43] for \( ggX \).

We perform a careful analysis of the different sources of the theoretical uncertainties and their correlations for \( \sigma_0^0, \sigma_0, \) and \( \sigma_2 \). The total theory covariance matrix for all parton types and bins under consideration is then obtained by adding the covariance matrices of all sources discussed below,

\[ C_{\text{th}} = C_{\text{FO}} + C_{\text{resum}} + C_{\varphi} + C_{\text{PDF}}. \]

(7)

For the perturbative uncertainties we follow the treatment developed in Refs. [27,32,44] and distinguish various independent sources. The first is an overall fixed-order yield uncertainty, \( C_{\text{FO}} \), which is fully correlated between all bins, and reproduces the usual fixed-order uncertainty for

the inclusive cross section. The resummation uncertainty, \( C_{\text{resum}} \), is induced by the binning cut and is correspondingly treated as a migration uncertainty that is fully anticonsorrelated between \( \sigma_0 \) and \( \sigma_0^{\text{nons}} \) and drops out of \( \sigma_2 \). The individual uncertainty contributions are estimated using profile scale variations [45,46] for the relevant resummation scales, as discussed in detail in Ref. [32]. Finally, we use a complex hard scale \( \mu_H = -im_X \) to resum large virtual QCD corrections. The corresponding resummation uncertainty, \( C_{\varphi} \), is estimated by varying the phase of \( \mu_H \) and corresponds to a yield uncertainty. The perturbative uncertainties are treated as fully correlated among all quark flavors and uncorrelated between quarks and gluons. We use the MMHT2014nnlo68cl [47] PDFs with the corresponding \( \alpha_s(m_Z) = 0.118 \) and 3-loop running. The parametric PDF uncertainties, \( C_{\text{PDF}} \), are constructed from the 25 independent eigenvectors of MMHT2014nnlo68cl. They are subdominant compared to the perturbative uncertainties.

III. INITIAL-STATE DISCRIMINATION

As a demonstration of the technique, we consider a hypothetical scalar resonance of mass \( m_X = 750 \) GeV produced in 13 TeV pp collisions. In Fig. 1 we show the ratio \( \sigma_0(p_{T}^{cut})/\sigma_2(p_{T}^{cut}) \) as a function of \( p_{T}^{cut} \) for different initial states \( i = u, c, b, g \) including all theoretical uncertainties, see Eq. (7). For all our results we use a jet radius of 0.4. The analogous results for \( d \) and \( s \) quarks are mostly indistinguishable from \( u \) quarks and are not shown. The split of the cross section into the 0-jet and \( \geq 1 \)-jet bins is clearly different for the different initial states, allowing one to distinguish light quarks (\( u,d,s \)), \( c \) quarks, \( b \) quarks, and gluons. The discrimination between \( b \) quarks and gluons is less good, in part due to the sizeable uncertainty in the gluon cross sections.

The optimal \( p_{T}^{cut} \) value for discriminating between different possible initial states depends on the true initial
state. Numerically, we observe only a mild sensitivity of the discrimination between initial states on the $p_{\text{cut}}^T$ value within the range $p_{\text{cut}}^T \in [25, 65]$ GeV. The $p_{\text{cut}}^T$ value can thus be chosen to optimize the experimental sensitivity with limited statistics. A roughly equal split of the cross section is achieved at $p_{\text{cut}}^T \approx 25$ GeV (65 GeV) for a light-quark (gluon) induced signal. In our subsequent analysis we use $p_{\text{cut}}^T = 40$ GeV, for which the cross section ratio is between 0.5 and 2 for any initial state.

Note that effects from hadronization and multiparton interactions, which are not included in our calculations, can affect the leading jet $p_T$ spectrum at small $p_T$. However, their effects partially compensate each other. We checked that the net effect in the cross section ratios we consider becomes negligible above $p_{\text{cut}}^T \gtrsim 20$ GeV.

To study the constraints on the Wilson coefficients $C_i$ from measuring $\sigma_0$ and $\sigma_{\geq 1}$, we minimize the $\chi^2$ function

$$\chi^2(C_i) = \sum_{m,n\in\{0,\geq 1\}} (\sigma_{m,\text{meas}} - \sigma_m)(C^{-1})_{mn}(\sigma_{n,\text{meas}} - \sigma_n),$$

where $\sigma_{m,\text{meas}}$ is the measured $pp \to X \to F$ cross section in bin $m$, $C$ is the sum of the experimental and theory covariance matrices, and $\sigma_m$ is the predicted cross section in bin $m$ in Eq. (5). In the narrow-width approximation and considering a single decay channel (e.g. $F = \gamma\gamma$), we only constrain $C_i \sqrt{B}$ [see Eq. (4)]. To render our results independent of the details of $F$, we define

$$C_i^{\text{incl}} \sqrt{B} = \sqrt{\sigma_{0,\text{meas}}/\sigma_{0}},$$

which is the value of $C_i \sqrt{B}$ for which the measured inclusive cross section is completely attributed to initial state $i$. By considering the ratios $|C_i/C_i^{\text{incl}}|$, our analysis only depends on the ratio $\sigma_0/\sigma_{\geq 1}$ but not the absolute cross sections or $B$. With more than one final state, the sum in Eq. (8) runs over the respective bins for all final states, and the dependence of the different branching ratios on the Wilson coefficients $C_i$ and the associated uncertainties need to be taken into account.

![Figure 2](image-url)

**FIG. 2.** $\Delta\chi^2 = 1$-contours for various scenarios. (a) gluon signal, (b) mixed gluon/$u$-quark signal, (c) $u$-quark signal, (d) $b$-quark signal, (e) $c$-quark signal, (f) gluon signal. The constraints from $\sigma_0$ and $\sigma_{\geq 1}$ are shown by the blue and green bands, respectively. The combined constraint from both are shown by the orange/yellow regions. The inner darker regions correspond to theory uncertainties only, while the full lighter bands include both theory and assumed experimental uncertainties.
EXPLOITING JET BINNING TO IDENTIFY THE ...

Figure 2 shows the constraints on $|C_i/C_i^{\text{incl}}|$ that can be achieved for various scenarios of the assumed true values of the $C_i$. For this purpose, we assume that the inclusive cross section is measured with a relative uncertainty of 20%, which can be realistically expected not very long after a discovery of a new state. We split the cross section into measured 0-jet and $\geq 1$-jet bins according to the theoretical predictions for the assumed signal, with the resulting $\sigma_0^{\text{meas}} / \sigma_1^{\text{meas}}$ given in each plot. The relative uncertainties, $\Delta\sigma_m$, on the measurements in the two bins are assumed to be uncorrelated and split according to $\Delta\sigma_0^{\text{meas}} / \Delta\sigma_1^{\text{meas}} = \sqrt{\sigma_0^{\text{meas}} / \sigma_0^{\text{meas}}}.$

In Fig. 2, contours of $\Delta^2 = 1$ only including the theoretical uncertainties are shown by the inner darker bands and combining theoretical and assumed experimental uncertainties by the outer lighter bands. The individual constraints from $\sigma_0$ and $\sigma_1$ are shown by the blue and green bands, respectively, while the combined constraint from both is shown in yellow/orange.

Figures 2(a)–2(c) illustrates the good discrimination between light quarks, here $u$, and gluons in the initial state, for a purely gluon-induced signal in (a), for a mixed signal with the cross section ratio equal to one in (b), and a purely $u$-quark induced signal in (c). Figures 2(d) and 2(e) demonstrate the good discrimination between $u$ and $b$ quarks for a $b$-quark signal, and between gluons and $c$ quarks for a $c$-quark signal, respectively. Only the discrimination between $b$ quarks and gluons, shown in Fig. 2(f) for a gluon signal, remains challenging due to the weaker separation already seen in Fig. 1.

We conclude that even with fairly large experimental uncertainties, as expected soon after a potential discovery, a clear separation between different initial states can be achieved for most scenarios. A combined fit to all coefficients will of course require more data, and will then also benefit from using several $p_T^{\text{cut}}$ values. We stress that thanks to the used resummation framework, the theoretical uncertainties and correlations can be robustly estimated and are not a limiting factor, and if necessary, could also be reduced further.

ACKNOWLEDGMENTS

This work was partially supported by the DFG Emmy-Noether Grant No. TA 867/1-1, the Office of Nuclear Physics of the U.S. Department of Energy under the Grant No. DE-SC0011090, the Collaborative Research Center SFB 676 of the DFG “Particles, Strings and the Early Universe,” the Simons Foundation through the Investigator Grant No. 327942, the Netherlands Organization for Scientific Research (NWO) through a VENI Grant, the PIER Helmholtz Graduate School, and by a Global MISTI Collaboration Grant from MIT.