Exploration of New Methods for Lattice QCD

by

Andres Rios Tascon

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

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2017

© Massachusetts Institute of Technology 2017. All rights reserved.

Author ........................................................................ Andrés Cózor Tascon
Department of Physics
May 12, 2017

Certified by ..........................................................
William Detmold
Assistant Professor of Physics
Thesis Supervisor

Accepted by ...........................................................
Nergis Mavalvala
Senior Thesis Coordinator, Department of Physics
Exploration of New Methods for Lattice QCD

by

Andres Rios Tascon

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

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2017

© Massachusetts Institute of Technology 2017. All rights reserved.
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.
Exploration of New Methods for Lattice QCD

by

Andres Rios Tascon

Submitted to the Department of Physics
on May 12, 2017, in partial fulfillment of the
requirements for the degree of
Bachelor of Science in Physics

Abstract

We explore two methods aimed at alleviating two difficulties in Lattice QCD: statistical noise and data storage. The first method intends to improve the signal-to-noise (S/N) ratio in three-point correlators, by extending previous work by Detmold and Endres. We test the method in the measurement of two observables: the nucleonic axial charge, and a matrix element computation related to the electromagnetic form factor of the rho meson. Only in the case of the rho we see a very slight improvement. We conclude that, in general, a case-by-case study would be needed to determine the effectiveness of the S/N optimization.

The second method that we study aims to improve data compression of gauge fields. It consists in generating a set of matrices distributed roughly uniformly along the group manifold, and constructing fine lattices around each of these matrices. We show that this compression can indeed provide better performance for SU(2) and SU(3) than the Lie algebra compression, but the improvement is not very significant. Nevertheless, we show that it is fairly close to the best compression one could hope to achieve with this type of method. We find that the compression procedure is very costly, which makes it currently impractical for machine precision compression. We conclude that studies must be done to determine if it can be improved by using additional information from the gauge fields.

Thesis Supervisor: William Detmold
Title: Assistant Professor of Physics
Acknowledgments

I am deeply grateful to Professor William Detmold, my research supervisor. I learned an incredible number of things working with him over the past year. I am also grateful to my academic advisor, Professor Hong Liu, and my previous research advisor, Professor Markus Klute.

I would also like to thank my parents and brother for their constant support throughout my undergraduate studies, and my friends at MIT and elsewhere.
# Contents

## 1 Introduction
- 1.1 Lattice Field Theory .................................................. 15
- 1.2 The Discretized Path Integral ...................................... 16
- 1.3 Monte Carlo Integral .................................................. 17
- 1.4 Lattice QCD ............................................................... 18
- 1.5 Correlation Functions .................................................. 19
- 1.6 Variational Method .................................................... 22
- 1.7 Lattice QCD Software .................................................. 24

## 2 Signal-to-Noise Optimization
- 2.1 S/N Optimization of Two-Point Correlators ....................... 25
  - 2.1.1 S/N Optimization with a Fixed Source ....................... 28
  - 2.1.2 S/N Optimization with Unconstrained Source and Sink .... 29
  - 2.1.3 Two-State Toy Model .............................................. 29
- 2.2 S/N Optimization of Three-Point Correlators ..................... 32
  - 2.2.1 Two-State Toy Model .............................................. 33
- 2.3 Applications: Hadronic Three-Point Functions .................... 35
  - 2.3.1 Nucleon Axial Charge ............................................ 36
  - 2.3.2 Rho Meson Form Factor .......................................... 38
- 2.4 Discussion and Conclusions ........................................... 40

## 3 Gauge Field Data Compression
- 3.1 Compression Method Description ..................................... 45
3.2 Theoretical performance ........................................ 47
3.3 Implementation ...................................................... 48
3.4 Application: SU(2) Field ............................................. 50
3.5 Application: SU(3) Field ............................................. 53
3.6 Discussion and Conclusions ......................................... 54

4 Conclusions ......................................................... 57
List of Figures

1-1 An example of an effective mass plot. At early times there is excited state contamination, while at late times the noise increases exponentially. 21

1-2 An example of the effectiveness of the variational method. Excited state contamination at early times is reduced, resulting in a larger plateau. 24

2-1 Comparison of the purely variational, variational plus S/N optimized, and purely S/N optimized correlators. 29

2-2 S/N landscape of two-point correlators of the two-state system with $a = 0.04$ and $b = c = 0.01$, using the parametrization from (2.21). The lower right plot shows a cross-sectional view of the other three plots at $\delta = 0$. 31

2-3 S/N landscape of three-point correlators of the two-state system with $a = 0.04$ and $b = c = 0.01$, using the parametrization from (2.21). The lower right plot shows a cross-sectional view of the other three plots at $\delta = 0$. 35

2-4 Comparison of the location of the S/N maximum for two- and three-point functions of the two-state system with $a = 0.04$ and $b = c = 0.01$, using the parametrization from (2.21). For large times (small $\Delta$) the locations of the maxima are similar. 35
2-5 Comparison of the variational method, variational plus S/N optimization, and S/N optimization for the nucleon axial charge for a sink separation of $t = 15$. The S/N optimization offers no improvement over the variational method in this case. 

2-6 Comparison of the results obtained from the fits for the nucleon axial charge for different time separations. The S/N optimization offers no improvement over the variational method in this case.

2-7 Same as Fig. 2-6, but without enforcing (anti)hermiticity. The S/N optimization for time separations $t = 15$ and $t = 18$ offers a slightly more precise result, but it is likely less accurate.

2-8 Comparison of the variational method, variational plus S/N optimization, and S/N optimization for the rho form factor for a sink separation of $t = 15$. The S/N optimization offers a very slight improvement.

2-9 Comparison of the results obtained from the fits for the rho form factor for different sink separation times. The S/N optimization offers only a very slight improvement for $t = 15$.

2-10 Same as Fig. 2-8, but without enforcing hermiticity. The S/N optimization significantly reduces the noise of individual data points, but the plateaus become less pronounced.

2-11 Same as Fig. 2-9, but without enforcing hermiticity. The S/N optimization offers practically no improvement overall, since the more subtle plateaus inhibit the improvement in the statistical noise.

3-1 Distribution of distances between random $SU(2)$ matrices and the closest element in $\Sigma = G120$.

3-2 Distribution of distances between products of $R_i$ (as in Eq. (3.5)) and the identity, in the case of $SU(2)$. Shown are $\epsilon = 0.06, 0.09, 0.12, 0.15$, with 16 bits.
3-3 Distance from random matrices to the closest matrix in the lattice. Increasing the number of bits reduces the average distance, as well as improving the behavior of the tail.  

3-4 Average and maximum distance between random $SU(2)$ matrices and the closest element in the lattice as a function of bits used for the fine lattice. The lines correspond to the best exponential fits.
List of Tables

3.1 Results from using 5 sets of matrices $R_i$ and 1000 random $SU(2)$ matrices. The increase in kurtosis and decrease in elapsed time at 22 bits is due to the switch to a different partitioning. .......................... 51

3.2 Comparison of the estimated number of bits required to describe a matrix with this method and other methods, as well as the best theoretical compression. ......................................................... 53

3.3 Same as Table 3.2, but for $SU(3)$. ......................................................... 54
Chapter 1

Introduction

1.1 Lattice Field Theory

Quantum Field Theories (QFTs) comprise the current theoretical framework for describing the elementary particles and interactions of nature, and also serve to describe composite particles and quasiparticles in condensed matter systems. Many observables, such as lepton-lepton scattering cross sections, can be found using perturbative techniques. However, many other observables can only be studied with nonperturbative calculations, which generally present a more formidable challenge. Significant progress in the nonperturbative regime started to be made with the introduction of lattice gauge theory by Wilson in 1974 [1]. The idea behind it was to discretize spacetime into a four-dimensional lattice, and formulate a field theory based on a discretized path integral. The theory was initially studied analytically, but a few years later Creutz, Jacobs and Rebbi performed the first numerical studies [2]. Since then, there has been enormous progress thanks to the rapid advancement of computers. This approach has mainly been used in the study of Quantum Chromodynamics (QCD), the theory describing quarks and gluons, since its complex nonlinear dynamics conceal many aspects of the theory. In this thesis we specifically focus on exploring improvements of techniques used for Lattice QCD.
1.2 The Discretized Path Integral

The quantization of a lattice field theory can be done using a discretized version of the path integral formalism. We explain the idea by first considering a nonrelativistic quantum system in one spatial dimension. The propagator for such a system can be written as

\[ \langle x_f, t_f | x_i, t_i \rangle = \langle x_f | e^{-\hat{H}(t_f-t_i)} | x_i \rangle = \int \mathcal{D}x(t)e^{-S[x]}, \]  

where \( \hat{H} \) is the Hamiltonian, \( S[x] = \int L[x]dt \) is the classical action, and \( \mathcal{D}x \) denotes the integration over all paths between \( (x_i, t_i) \) and \( (x_f, t_f) \). Note that there are factors of \( i \) missing from the usual expressions. This is because a Wick rotation is performed to obtain an Euclidean signature. This is crucial for evaluating the integral numerically, as will be mentioned below.

A path \( x(t) \) can, in principle, be arbitrarily complicated, which would be impossible to handle by a computer. But if time is discretized into \( N \) slices, then the path \( x(t) \) becomes a vector of the positions at times \( t_n = an \) for a spacing \( a = (t_f - t_i)/N \),

\[ x = (x(t_0), x(t_1), \cdots, x(t_N)) = (x_0, x_1, \cdots, x_N), \]  

where we introduced the notation \( x_n = x(t_n) \), and where \( x_0 = x_i \) and \( x_N = x_f \). We refer to this vector as a "configuration." Additionally, by replacing derivatives with finite difference approximations one can write an action \( S_{\text{lat}} \) that only depends on this vector. Thus, one obtains a formulation of the theory that is suitable for computers. The propagator becomes

\[ \langle x_f, t_f | x_i, t_i \rangle = A \int dx_1dx_2\cdots dx_{N-1}e^{-S_{\text{lat}}([x_0,x_1,\cdots,x_N])}. \]  

where \( A \) is some constant. Note that we do not integrate over \( x_0 \) or \( x_N \) because the endpoints are fixed.

In general, the expectation value of an operator \( \hat{\Gamma} \) is given in the path integral formulation by

\[ \langle \hat{\Gamma} \rangle = \frac{\int \mathcal{D}x(t)\hat{\Gamma}[x]e^{-S[x]}}{\int \mathcal{D}x(t)e^{-S[x]}} \]  

(1.4)
where $\Gamma$ is interpreted as a functional on the right side. Hence, in a lattice with $N$ time slices and periodic boundary conditions (i.e. $x_0 = x_N$), the expectation value is given by

$$\langle \hat{\Gamma} \rangle = A \int dx_0 dx_2 \cdots dx_{N-1} \Gamma[(x_0, x_1, \cdots, x_{N-1})] e^{-S_{\text{int}}[(x_0, x_1, \cdots, x_{N-1})]}$$

(1.5)

where $A$ is some constant.

### 1.3 Monte Carlo Integral

Although it is possible to compute expectation values with Eq. (1.5) for simple systems, it becomes extremely impractical for complex systems, such as fields, where the number of integrals would typically be above $10^6$. Alternatively, since the theory was formulated using Euclidean time, one can recognize that Eq. (1.5) has the form of a weighted average with weight $e^{-S[x]}$. Hence, a Monte Carlo procedure can be used to estimate expectation values. To do this, one generates a large number $N_{\text{cf}}$ of configurations,

$$x^{(\alpha)} = (x_0^{(\alpha)}, x_1^{(\alpha)}, \cdots, x_N^{(\alpha)}), \quad \alpha = 1, 2, \cdots, N_{\text{cf}},$$

(1.6)

such that the probability for generating a configuration follows $P[x^{(\alpha)}] \propto e^{-S[x^{(\alpha)}]}$. Then, the expectation value can be approximated as

$$\langle \hat{\Gamma} \rangle = \frac{1}{N_{\text{cf}}} \sum_{\alpha=1}^{N_{\text{cf}}} \Gamma[x^{(\alpha)}].$$

(1.7)

There are various procedures that can be used to generate configurations that follow the desired probability. The simplest one is the Metropolis algorithm [3], which consists in progressively updating previous configurations in a way that resembles the quantum fluctuations of the theory. However, it has the disadvantage of the resulting configurations being very correlated. There are other procedures that reduce the correlations between configurations (see e.g. [4, 5, 6]), but as long as they are used carefully they all produce equally good configurations with the desired probability.
1.4 Lattice QCD

When the system in consideration is a field, the position $x(t)$ above is replaced by a field variable $\phi(x)$. Analogously, the coordinates $x^\mu$ are now discretized, and the integrals are over $d\phi(x)$. We assume an isotropic lattice with spacing $a$. There are many additional subtleties when dealing with fields. For instance, the bare parameters must be renormalized due to the lattice spacing serving as a momentum cutoff, and the discretization of the action must be done carefully to not introduce nonphysical effects. For an overview of the intricacies see Ref. [7].

In the specific case of QCD there is a gauge field $A_\mu$ for the gluons, with $SU(3)$ as the gauge group, a fermionic field $\Psi$ for the quarks, and a coupling constant $g$. The coupling is commonly expressed as a $\beta$ parameter defined by $\beta \equiv 6/g^2$. In the lattice formulation, fermionic fields can be thought of as living in the sites (nodes of the lattice), but gauge fields are not as straightforward. It turns out that, to preserve gauge invariance of the fermionic action and construct a lattice gauge action, the gauge field must be replaced by a field $U_\mu$. In the continuum theory, it corresponds to the path-ordered exponential integral of $A_\mu$ (i.e. Wilson line) from one site to the next one,

$$U_\mu(x) = \mathcal{P} \exp \left[ ig \int_x^{x+a\vec{a}} A_\nu(y) dy^\nu \right].$$

Thus, the gauge field can be though of as living in the links of the lattice (lines joining adjacent sites).

One of the main developments in Wilson’s paper [1] is the introduction of a lattice gauge action. It is constructed using the trace of the ordered product of links around the elementary squares of the lattice (referred to as plaquettes). The rest of the QCD action is constituted by the discretized fermionic action due to the quarks.

Thanks to the anticommuting property of fermionic variables, the fermionic part of the action can be integrated out. Hence, expectation values can be calculated by averaging over gauge configurations weighted by some effective action. Thus, we use an analogous version of Eq. (1.5) to find an expectation value as an average over gauge configurations.
For the above reason, computations in Lattice QCD require the storage of numerous gauge configurations. Hence, it is desirable to have an efficient method of storing them. This will be our focus in Chapter 3.

1.5 Correlation Functions

Euclidean two-point correlators are defined as

\[ \langle O_2(t)O_1(0) \rangle_T = \frac{1}{Z_T} \text{Tr} \left[ e^{-(T-t)\hat{H}} \hat{O}_2 e^{-t\hat{H}} \hat{O}_1 \right] \quad \text{with} \quad Z_T = \text{Tr} \left[ e^{-T\hat{H}} \right], \quad (1.9) \]

where we interpret \( T \) as a formal maximal time. Depending on the choice of boundary conditions, one could also interpret \( T \) as an inverse temperature. We can choose an orthonormal basis \( \{ |n\rangle \} \) of eigenstates of \( \hat{H} \) with eigenvalues \( E_n \) such that \( E_n \leq E_m \) if \( n < m \). Thus, can rewrite the expression above as

\[ \langle O_2(t)O_1(0) \rangle_T = \frac{1}{Z_T} \sum_{m,n} e^{-(T-t)E_n} \langle m|\hat{O}_2|n\rangle e^{-tE_m} \langle n|\hat{O}_1|m\rangle \]

with \( Z_T = \sum_n e^{-TE_n} \). \quad (1.10)

Defining \( \Delta E_n = E_n - E_0 \) we write

\[ \langle O_2(t)O_1(0) \rangle_T = \frac{\sum_{m,n} \langle m|\hat{O}_2|n\rangle \langle n|\hat{O}_1|m\rangle e^{-t\Delta E_n} e^{-(T-t)\Delta E_m}}{\sum_m e^{-T\Delta E_m}}. \quad (1.11) \]

In the limit \( T \to \infty \) only the terms with \( m = 0 \) will remain. Thus, in this limit we recover the usual vacuum expectation value (assuming that the vacuum is unique and that \( E_1 > 0 \)). We will always assume that the lattice is large enough so that this limit holds. For a more detailed discussion see Ref. [7].

In the lattice formulation, \( t \) is taken to be a dimensionless integer which enumerates the time slice, and other quantities (e.g. the energy \( E \)) absorb a factor of the spacing \( a \). Thus, many of the results obtained will be in lattice units, meaning that they must be multiplied by an appropriate factor of \( a \) to obtain the usual units.
The process shown above can be generalized for general expectation values. In this thesis we will only consider two- and three-point correlators. Let us first discuss two-point functions. Let $O^\dagger$ be some creation operator, and $O$ its conjugate (annihilation) operator. We refer to these operators as “interpolators.” We have

$$
\langle O(t)O^\dagger(0) \rangle = \sum_n \langle \Omega|\hat{O}|n\rangle \langle n|\hat{O}^\dagger|\Omega\rangle e^{-tE_n} = Ae^{-tE_H}(1 + \mathcal{O}(e^{-t\Delta E})),
$$

(1.12)

where $|\Omega\rangle$ is the vacuum state, $A$ is some constant, and $E_H$ is the energy of the lowest energy state with the same quantum numbers as $O^\dagger$. Note that for large $t$ the two-point function behaves as $e^{-tE_H}$, which we can use to find $E_H$.

It is convenient to consider states with a fixed momentum. We can perform a Fourier transform to project the sink to momentum $\vec{p}$,

$$
\tilde{O}(\vec{p}, t) = \frac{1}{\sqrt{|\Lambda_3|}} \sum_{\vec{n} \in \Lambda_3} O(\vec{n}, t)e^{-i\vec{n}\cdot\vec{p}},
$$

(1.13)

where $\Lambda_3 = \{(n_1, n_2, n_3)|n_i = 0, 1, 2, \cdots, N - 1\}$ and $|\Lambda_3|$ is the spatial volume. It is also customary to place the creation operator at the origin.

It is usually the case that the state of interest is at rest, so the mass (or energy) can be found using

$$
C(t) \equiv \langle \tilde{O}(0, t)\tilde{O}^\dagger(0, 0) \rangle = \sum_n \langle \Omega|\hat{O}|n\rangle \langle n|\hat{O}^\dagger|\Omega\rangle e^{-tE_n} = Ae^{tE_H}(1 + \mathcal{O}(e^{-t\Delta E})),
$$

(1.14)

where now $E_H$ can be interpreted as the mass of the lowest energy state of total momentum zero (if it is a single-particle state). More generally, we can define an effective mass as

$$
m_{\text{eff}}(t + \frac{1}{2}) = \ln \left( \frac{C(t)}{C(t + 1)} \right),
$$

(1.15)
for baryonic states, and as

$$\frac{\cosh(m_{\text{eff}}(t - N/2))}{\cosh(m_{\text{eff}}(t + 1 - N/2))} = \frac{C(t)}{C(t + 1)},$$

(1.17)

(and equivalently with sinh) for mesonic states. A typical effective mass plot is shown in Fig. 1-1. The mass can be extracted by fitting the plateau region. Since time slices are correlated, the fit must be performed taking into account the correlation between data points. There are two important aspects to notice from the figure, which are general to all effective mass plots. Firstly, at early times there is excited state contamination, which is expected from Eq. (1.15). And secondly, at late times the statistical noise grows very rapidly. The exponential growth of the noise will be our focus in Chapter 2.

![Figure 1-1: An example of an effective mass plot. At early times there is excited state contamination, while at late times the noise increases exponentially.](image)

Let us now consider three-point functions. We will focus on those with the form

$$\langle \hat{O}_2(\vec{p}, t)\hat{\chi}(\vec{q}, \tau)\hat{O}_1(\vec{0}, 0) \rangle,$$

where $\chi$ is a current operator (e.g. $\bar{\psi}\gamma^\mu\psi$), $\vec{p}$ is the momentum of the sink, $\vec{q}$ is the momentum transfer, and we define $r \equiv p - q$ as the momentum of the source. These correlators are used in the study of hadron structure to compute form factors. Using the same arguments as above we can find that for $1 \ll \tau \ll t$ we have

$$\langle \hat{O}_2(\vec{p}, t)\hat{\chi}(\vec{q}, \tau)\hat{O}_1(\vec{0}, 0) \rangle = \langle \Omega|\hat{O}_2|\vec{p}\rangle e^{-E_p(t - \tau)}\langle \vec{p}|\hat{\chi}|\vec{r}\rangle e^{-E_r\tau}\langle \vec{r}||\hat{O}_1|\Omega \rangle,$$

(1.18)
where $|\vec{p}\rangle$ (resp. $|\vec{r}\rangle$) is the lowest energy state with the quantum numbers of $O_2$ (resp. $O_1$) and its energy is $E_p$ (resp. $E_r$). We are specifically interested in the matrix element $\langle \vec{p}|\chi|\vec{r}\rangle$ since it corresponds to the coupling of the current with some specific particle state. For instance, if the operators $O_1$ and $O_2$ had the quantum numbers of the rho meson, then this matrix element would be $\langle \rho(p)|\chi|\rho(r)\rangle$. If the current is $\chi^\mu = \frac{2}{3}\bar{u}\gamma^\mu u - \frac{1}{3}\bar{d}\gamma^\mu d$, the matrix element is proportional to the electromagnetic form factor of the rho, $F^{(\rho)}(q^2)$, which can be used to find its charge radius. The matrix element $\langle \vec{p}|\chi|\vec{r}\rangle$ can be obtained by computing $R$, defined as follows

$$R(\vec{p}, \vec{q}, \tau, t) \equiv \frac{\langle \tilde{O}_2(\vec{p}, t)\tilde{x}(\vec{q}, \tau) O_1(0, 0) \rangle}{\langle \tilde{O}_2(\vec{p}, t) O_1(0, 0) \rangle} \times \frac{\langle \tilde{O}_2(\vec{q}, t)O_1(0, 0)\rangle}{\sqrt{\langle \tilde{O}_2(\vec{q}, t) O_1(0, 0) \rangle}} \frac{\langle \tilde{O}_2(\vec{r}, t-\tau) O_1(0, 0) \rangle}{\langle \tilde{O}_2(\vec{r}, t) O_1(0, 0) \rangle}.$$  \hspace{1cm} (1.19)

Note that for $\vec{q} = 0$ the terms in the square root cancel out.

## 1.6 Variational Method

Measuring observables in Lattice QCD usually requires fitting a plateau region, as it is the case for effective masses. Hence, it is desirable to make the plateau region as large as possible. This can be achieved by reducing the excited state contamination at early times. One possible way of doing this is by using smeared sources, which tend to better approximate the wave function of the state in question. Another method, which has proven to be highly successful, is the Variational Method [8, 9, 10]. It works by constructing a matrix of correlators $C_{ij}$ using different source and sink operators (with the same quantum numbers),

$$C_{ij}(t) = \langle O_i^\dagger(t) O_j(0) \rangle.$$  \hspace{1cm} (1.20)

We will suppress the tilde and some arguments for notational convenience, but the sink should always be thought as projected to some definite momentum and the source as set at the origin. By inserting a complete set of energy eigenstates (with energies
ordered as above), we may write

\[ C_{ij}(t) = \langle \Omega | O'_i e^{-\hat{H}t} O^\dagger_j | \Omega \rangle = \sum_n Z'_i n Z^*_j n e^{-E_n t}, \tag{1.21} \]

where \( Z'_i n = \langle \Omega | \hat{O}'_i | n \rangle \) and \( Z_j n = \langle \Omega | \hat{O}_j | n \rangle \) are overlap factors. Henceforth, we will adopt the convention that primed quantities are associated with sinks and unprimed quantities with sources (except for the current operator).

A general correlator can be constructed as a linear combination of sources and sinks. In matrix notation, we write \( c(\psi', \psi) = \psi'^\dagger C \psi \), where \( \psi \) and \( \psi' \) are source and sink vectors respectively. We normalize the vectors so that \( \psi^\dagger \psi = 1 \) and \( \psi'^\dagger \psi' = 1 \).

Usually, the primed and unprimed interpolators are chosen from the same set \( \{ O_1, \ldots, O_N \} \), so that \( C \) is a \( N \times N \) hermitian matrix. It has been shown that by diagonalizing this matrix one can better isolate the lowest energy states better. The eigenvalues can be shown to behave as

\[ \lambda^{(k)}(t) \propto e^{-t E_k} (1 + \mathcal{O}(e^{-t \Delta E_k})), \tag{1.22} \]

where the energies are arranged in ascending order, and \( \Delta E_k \) is the distance from \( E_k \) to nearby energy levels. The eigenvalues and eigenvectors can be found with the generalized eigenvalue problem

\[ Q(t_0)^{-1} C(t) Q^\dagger(t_0)^{-1} \vec{v} = \lambda(t) \vec{v}, \tag{1.23} \]

where \( Q(t_0)Q^\dagger(t_0) = C(t_0) \) and \( t_0 < t \) is some early time.

For a given \( \lambda^{(k)} \), the corresponding eigenvector \( \psi \) (which is equal to \( \psi' \) if the matrix is hermitian) corresponds to a linear combination of interpolators that gives the best overlap with the state of energy \( E_k \). Thus, the contamination from other states is decreased. The result of applying the variational method using five sources and sinks is shown in Fig. 1-2. It is clear from the figure that excited state contamination is reduced at early times. Thus, a larger plateau is present, which results in a more precise determination of the mass.
Figure 1-2: An example of the effectiveness of the variational method. Excited state contamination at early times is reduced, resulting in a larger plateau.

In Chapter 2 we will compare the variational method with an alternative approach to improve precision, which aims to reduce the noise at late times instead of (or as well as) minimizing excited state contamination.

1.7 Lattice QCD Software

There are several freely available software packages designed to generate gauge configuration and compute quark propagators. We used Chroma[11] with GPU acceleration by QUDA[12], which are maintained by the USQCD collaboration and M. Clark, respectively. They were used in Chapter 2 to generate gauge configurations and compute two- and three-point correlators.
Chapter 2

Signal-to-Noise Optimization

Measurements in Lattice QCD suffer from two main complications that limit their precision: excited state contamination and statistical noise. Techniques such as the variational method reduce the effects of the former complication, but not much has been done to address the latter since it mainly depends on the finite statistics available for the computation. In 2014, Detmold and Endres introduced a method to alleviate the exponential reduction of the signal-to-noise (S/N) ratio at late times in two-point functions [13]. In this chapter we explore an extension of this technique to three-point functions, which could be useful in the study of hadron structure.

We start by reviewing how the S/N optimization was derived in Ref. [13].

2.1 S/N Optimization of Two-Point Correlators

The S/N optimization follows the same idea as the variational method, discussed in Sec. 1.6, of constructing a matrix of two-point correlators with different source and sink interpolators (with the same quantum numbers). However, instead of finding source and sink vectors that maximize the overlap with a specific state, it constructs ones that optimize the S/N ratio. Recall that unprimed quantities are associated with the sources and primed quantities with the sinks. As in Sec. 1.6, we consider a
$N' \times N$ matrix $C$ of two-point correlators,

$$C_{ij}(t) = \sum_{n} Z'_{in} Z^*_{jn} e^{-E_{nt}},$$  \hspace{1cm} (2.1)

where $Z'_{in} = \langle \Omega | \hat{O} | n \rangle$ and $Z_{jn} = \langle \Omega | \hat{O}_j | n \rangle$. As before, we assume that the sink operator is projected to a specific momentum and that the source operator is located at the origin.

Following an argument by Lepage [14], when calculating a two-point correlator $\langle \Omega | \hat{O}(t) \hat{O}^\dagger(0) | \Omega \rangle$ where the quark degrees of freedom have been integrated out and are expressed as quark propagators, the main noise contribution will be due to states produced by $\hat{O} \hat{O}^\dagger$ (i.e. in the variance correlator $\langle \langle \hat{O} \hat{O}^\dagger \rangle^\dagger(t) (\hat{O} \hat{O}^\dagger)(0) \rangle$). More specifically, the associated uncertainty on the estimate is given by the standard deviation of the correlator distribution $\sigma_{c}(\psi', \psi)/\sqrt{N}$, where $N$ is the ensemble size and

$$\sigma_{c}^2(\psi', \psi) = \sigma^2(\psi', \psi) - |\psi'^\dagger C \psi|^2,$$  \hspace{1cm} (2.2)

is the variance, with

$$\sigma^2(\psi', \psi) = (\psi' \otimes \psi'^*) \Sigma^2(\psi \otimes \psi^*), \quad \Sigma^2 \equiv \langle C \otimes C^* \rangle.$$  \hspace{1cm} (2.3)

We also define the following matrices for future use

$$\sigma_{\psi}^2 = \langle C \psi \psi^\dagger C^\dagger \rangle, \quad \sigma_{\psi'}^2 = \langle C'^\dagger \psi' \psi'^\dagger C \rangle,$$  \hspace{1cm} (2.4)

and note that they satisfy the relation

$$\sigma^2(\psi', \psi) = \psi'^\dagger \sigma_{\psi}^2 \psi' = \psi'^\dagger \sigma_{\psi'}^2 \psi.$$  \hspace{1cm} (2.5)

The “noise correlator” $\Sigma^2$ can be expanded as

$$\Sigma^2_{ik;jl}(t) = \sum_{n} \tilde{Z}'_{ik,n} \tilde{Z}_{jl,n}^* e^{-E_{nt}},$$  \hspace{1cm} (2.6)
where $\tilde{Z}_{ik,n} = \langle \Omega | \hat{O}_k \hat{O}_i^\dagger | n \rangle$ and $\tilde{Z}_{jl,n} = \langle \Omega | \hat{O}_j \hat{O}_l^\dagger | n \rangle$ are overlap factors. The states $| n \rangle$ are the "noise states" with the quantum numbers of $\hat{O}_j \hat{O}_l^\dagger$, and energies $\tilde{E}_n$ (ordered such that $\tilde{E}_n \leq \tilde{E}_m$ for $n < m$). The indices $i$ and $k$ run from 1, $\cdots$, $N'$, and $j$ and $l$ run from 1, $\cdots$, $N$.

The signal-to-noise ratio is thus given by

$$\theta_c(\psi', \psi) = \sqrt{N} \frac{|\psi'^\dagger C \psi|}{\sigma_c(\psi', \psi)},$$

or equivalently, by

$$\theta_c(\psi', \psi) = \sqrt{N} \left( \frac{1}{\theta^2(\psi', \psi)} - 1 \right)^{-1/2},$$

where

$$\theta(\psi', \psi) = \frac{|\psi'^\dagger C \psi|}{\sigma(\psi', \psi)}.$$

Hereafter, we will ignore the $\sqrt{N}$ factor. Note that since $\sigma^2_c$ is nonnegative, then $0 \leq \theta(\psi', \psi) \leq 1$ and $\theta_c(\psi', \psi)$ is a monotonically increasing function of $\theta(\psi', \psi)$. We are interested in the late-time behavior, when $\theta(\psi', \psi) \ll 1$, in which case we have $\theta_c = \theta$ to leading order.

We now consider a correlator constructed from a fixed source and sink vector. The late-time behavior of such correlator is given by

$$\psi'^\dagger C \psi \sim \psi'^\dagger Z_{0}^\dagger \sigma_{0} \psi e^{-E_0 t},$$

up to corrections of order $\Delta = e^{-(E_1 - E_0)t}$, where $E_0$ (resp. $E_1$) is the energy of the ground (resp. first excited) state with the quantum numbers of $O_i^\dagger$ and $O_j$. The late-time behavior of the variance is given by

$$\sigma^2(\psi', \psi) \sim (\psi'^\dagger \tilde{Z}_0^\dagger \psi')(\psi'^\dagger \tilde{Z}_0^\dagger \psi) e^{-\tilde{E}_0 t},$$

up to corrections of order $\tilde{\Delta} = e^{-(\tilde{E}_1 - \tilde{E}_0)t}$, where $\tilde{E}_0$ (resp. $\tilde{E}_1$) is the energy of the ground (resp. first excited) noise state with the quantum numbers of $O_i^\dagger O_k^\dagger$ and $O_j O_l^\dagger$. Since the variance $\sigma^2$ is positive-definite for all $\psi, \psi'$, then $\tilde{Z}_0, \tilde{Z}_0'$ are positive-definite.
when regarded as a two-index matrices. Using Eqs. (2.10) and (2.11) with Eq. (2.9) we see that the S/N ratio has the following late-time behavior

\[ \theta(\psi', \psi) \sim \frac{|\psi'^\dagger Z_0| |Z_0^\dagger \psi|}{\sqrt{\psi'^\dagger \tilde{Z}_0' \psi'} \sqrt{\psi^\dagger \tilde{Z}_0 \psi}} e^{-(E_0-\frac{1}{2}E_0)t}. \]  

(2.12)

Thus, we see that the S/N ratio decreases exponentially at late times. It can be show that this is maximized when \( \psi'_0 \propto (\tilde{Z}_0')^{-1}Z_0' \) and \( \psi_0 \propto (\tilde{Z}_0)^{-1}Z_0 \), with a maximum of

\[ \theta(\psi'_0, \psi_0) \sim \sqrt{Z_0'^{\dagger} (\tilde{Z}_0')^{-1}Z_0'} \sqrt{Z_0^{\dagger} (\tilde{Z}_0)^{-1}Z_0} e^{-(E_0-\frac{1}{2}E_0)t}. \]  

(2.13)

Furthermore, the maximum is unique. For details see Ref. [13].

Let us now discuss the procedure used to maximize the S/N in the two cases that will be relevant for this thesis (for other cases, refer to the original paper).

### 2.1.1 S/N Optimization with a Fixed Source

In Sec. 1.6 we discussed how the variational method is very effective at reducing excited state contamination, thus allowing for more precise measurements. It would be desirable to maintain the reduction in excited state contamination while also improving the S/N ratio. This can be done by fixing the source vector \( \psi \) to the one obtained with the variational method, and optimize the S/N with the sink vector \( \psi' \). For a given \( \psi \), the S/N optimized sink is given by

\[ \psi'_0 = A'_0(\psi) \sigma^{-2}_\psi C \psi, \]  

(2.14)

where \( A'_0(\psi)^{-2} \equiv \psi^\dagger C^\dagger \sigma^{-4}_\psi C \psi \), and \( \sigma_\psi \) is defined in (2.4).

Note, however, that since the source and sink vectors are different then \( \psi'_0 C \psi \) is no longer positive definite. Thus, it may deviate from the true expectation value or contain an imaginary part. For a sufficiently large basis and a well chosen source vector this effect should be small. A comparison of the variational method and the S/N optimized sink is shown in Fig. 2-1.
2.1.2 S/N Optimization with Unconstrained Source and Sink

In some cases it may be advantageous to improve the S/N ratio at the expense of introducing excited state contamination at early times. In order to find the best source and sink vectors an iterative procedure can be used. Starting with arbitrary vectors $\psi'_0$ and $\psi_0$ one defines

$$
\psi'^{[n+1]} = A'_0(\psi^{[n]})\sigma_{\psi^{[n]}}^{-2}C\psi^{[n]}, \quad \psi^{[n+1]} = A_0(\psi'^{[n]})\sigma_{\psi'^{[n]}}^{-2}C'^{\dagger}\psi'^{[n]},
$$

with normalization factors given by

$$
A'_0(\psi)^{-2} = \psi'^{\dagger}C'^{\dagger}\sigma_{\psi'}^{-4}C\psi, \quad A_0(\psi')^{-2} = \psi'^{\dagger}C\sigma_{\psi'}^{-4}C'^{\dagger}\psi'.
$$

If the procedure converges, which must be confirmed explicitly, the source and sink vectors will converge to those that maximize the S/N ratio. It can easily be checked that if the correlator matrix is hermitian the two vectors will be equal. A comparison of this technique and the variational method is shown in Fig. 2-1.

![Figure 2-1: Comparison of the purely variational, variational plus S/N optimized, and purely S/N optimized correlators.](image)

2.1.3 Two-State Toy Model

Detmold and Endres used a two-state system as a toy model to show the basic features and capabilities of the S/N optimization [13]. We will review their example and use
it to draw comparisons between the S/N landscape of two- and three-point functions.

A general $2 \times 2$ matrix of correlation functions may be written as

$$C \propto Z_0 Z_0^\dagger + \Delta Z_1 Z_1^\dagger,$$  \hspace{1cm} (2.17)

where $\Delta = e^{-(E_1 - E_0)t}$, expressed in the basis where

$$Z_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad Z_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$  \hspace{1cm} (2.18)

Note that if $Z_0$ and $Z'_0$ are different, one can make them equal by performing a change of basis, provided that the operator basis is complete, which we assume here. Similarly, the noise correlator is given by

$$\Sigma^2 \propto \tilde{Z}_0 \tilde{Z}_0^\dagger + \mathcal{O}(\tilde{\Delta}),$$  \hspace{1cm} (2.19)

where $\tilde{\Delta} = e^{-(\tilde{E}_1 - \tilde{E}_0)t}$. Since $\tilde{Z}_0$ is a positive-definite Hermitian matrix, it has the form

$$\tilde{Z}_0 = \begin{pmatrix} a & b \\ b^* & c \end{pmatrix},$$  \hspace{1cm} (2.20)

with $a > 0$, $c > 0$ and $ac > |b|^2$. It is also important to keep in mind that the signal and noise states usually have different quantum numbers, with the only exception being correlators with vacuum quantum numbers. Thus $\Delta$ and $\tilde{\Delta}$ generally differ.

We will study the S/N landscape of this system in the case when the source and sink vectors are equal, $\psi = \psi'$. These vectors can be parametrized by

$$\psi(\omega, \delta) = \begin{pmatrix} \cos \omega \\ \sin \omega e^{i\delta} \end{pmatrix},$$  \hspace{1cm} (2.21)

up to an irrelevant overall phase. The variables in this parametrization are in the ranges $\omega \in [0, \pi)$ and $\delta \in [-\pi/2, \pi/2)$. Apart from the dependence on these variables, we also consider the dependence of the S/N on $t$, which is implicit in $\Delta$.  

30
The signal-to-noise ratio for this model with the above parametrization, \( \theta_*(\omega, \delta) = \theta(\psi(\omega, \delta), \psi(\omega, \delta)) \), has the form

\[
\theta_*(\omega, \delta) = \theta_*(0, 0) \frac{|1 + 2\Delta \tan^2 \omega|}{1 + \frac{\xi}{a} \tan^2 \omega + 2\frac{|b|}{c} \tan \omega \cos(\arg(b) + \delta)},
\]

up to relative corrections of order \( \hat{\Delta} \). It has a global maximum \( \theta_*(\omega_*, \delta_*) = R_* \theta_*(0, 0) \) at a point \((\omega_*, \delta_*)\) satisfying \( \arg(b) + \delta_* = \pi \) and

\[
\frac{c}{a} = \frac{R_* - 1 + \Delta \tan^2 \omega_*}{R_* \tan^2 \omega_*}, \quad \frac{|b|}{a} = \frac{R* - 1}{R_* \tan \omega_*}.
\]

As a concrete example we consider the case when \( a = 0.04 \) and \( b = c = 0.01 \). Fig. 2-2 shows the S/N landscape using the parametrization from Eq. (2.21). Generally \( \theta_* \) is not symmetric about \( \delta = 0 \). This is the case here because the parameter \( b \) was chosen to be real. We will discuss further aspects of the plots when we consider the three-point functions of the same simple system in Sec. 2.2.1.

Figure 2-2: S/N landscape of two-point correlators of the two-state system with \( a = 0.04 \) and \( b = c = 0.01 \), using the parametrization from (2.21). The lower right plot shows a cross-sectional view of the other three plots at \( \delta = 0 \).
2.2 S/N Optimization of Three-Point Correlators

We now follow the same procedure as above, but for three-point correlators. Let \( C \) be a \( N' \times N \) matrix of three-point correlation functions in Euclidean spacetime constructed by using the interpolating operators \( \hat{O}'(t) \) and \( \hat{O}_j(0) \) (all with the same quantum numbers), and a current operator \( \hat{\chi}(\tau) \) at some intermediate time \( \tau \). Its matrix elements are given by

\[
C_{ij}(\tau, t) = \langle \Omega | \hat{O}'(p', t) \hat{\chi}(q, \tau) \hat{O}_j(0) | \Omega \rangle. \tag{2.24}
\]

This time we cannot suppress the momentum arguments since there are two relevant momenta: the sink momentum \( p' \) and the momentum transfer \( q \) (the source is left unprojected, located at the origin, but by momentum conservation \( p = p' - q \)). By inserting complete sets of energy eigenstates, as before, we obtain

\[
C_{ij}(\tau, t) = \sum_{n,m} Z'_{in,p'} e^{-(t-\tau)E_n,p'} \chi_{nm,q} e^{-\tau E_m,q} Z^*_{jm,q}, \tag{2.25}
\]

where \( Z'_{in,p'} = \langle \Omega | \hat{O}'(p') | n \rangle \) and \( Z_{jm,q} = \langle \Omega | \hat{O}_j(p) | m \rangle \) are overlap factors, \( E_{n,p} \) is the \( n \)th eigenstate with momentum \( p \), and \( \chi_{nm,q} = \langle n | \hat{\chi}(q) | m \rangle \) is the matrix element of the current operator with momentum transfer \( q \).

Following the same procedure as above we find that the noise correlator \( \Sigma^2 \) is given by

\[
\Sigma^2_{ik;jl}(\tau, t) = \sum_{n,m} Z'_{ik,n,p'} e^{-\frac{1}{2}\tilde{E}_n,p'}(t-\tau) (\chi^\dagger)_{nm,q} e^{-\tau \tilde{E}_m,q} Z^*_{jl,m,q}, \tag{2.26}
\]

where the overlap factors and energies follow the same convention as before. The final result turns out to be very resemblant of the previous one,
for $1 \ll \tau \ll t$. Thus, the procedures described in Sec. 2.1.1 and Sec. 2.1.2 should apply equally well to these correlators. Nevertheless, since in this case we take two limits, there can be more important subleading terms.

Although the expression for the S/N ratio is similar to the one for two-point functions, there are important things to note in this case. Mainly, there is much more freedom in three-point functions. In particular, the noise will not only be determined by the states, but also by the current operator. Even for the same initial and final states, different current operators can result in largely different levels of noise. A clear example is the case of the electromagnetic current (i.e. $\bar{\psi} \gamma^\mu \psi$), where the noise tends to be low. For zero momentum transfer, which corresponds to measuring the charge of the particle, there is a very clear plateau and almost no noise. In contrast, the axial current (i.e. $\bar{\psi} \gamma^\mu \gamma^5 \psi$) has relatively large levels of noise, even at zero momentum transfer. This is the reason why in Sec. 2.3 we examine the axial charge, and not the electromagnetic one. Furthermore, having a nonzero momentum transfer can result in initial and final states having different levels of noise and excited state contamination. We see this effect in Sec. 2.3 for the rho meson, where the plot of $R$ is very asymmetric.

### 2.2.1 Two-State Toy Model

We again consider a simple two-state system as a toy model. We will assume that the momentum transfer is zero, $\vec{q} = 0$, so that we can simplify the notation by omitting the momentum subscripts. Nevertheless, the following arguments can be made for any momentum transfer. We will also assume that $\chi$ is real and symmetric, that $0 \ll \tau \ll t$ and that $O(e^{-E_n(t-\tau)}) \sim O(e^{-E_n \tau})$. A general $2 \times 2$ matrix of three-point correlation functions may thus be given by

$$C \propto Z_0 Z_0^\dagger + \frac{\chi_{01}}{\chi_{00}} (Z_0 Z_0^\dagger + Z_1 Z_1^\dagger) + \Delta^2 \frac{\chi_{11}}{\chi_{00}} Z_1 Z_1^\dagger,$$  \hspace{1cm} (2.28)
where \( \Delta = e^{-(E_1-E_0)r} \sim e^{-(E_1-E_0)(t-r)} \), expressed in the basis where \( Z_0 = (1) \) and \( Z_1 = (i) \). Similarly, the noise correlator is given by

\[
\Sigma^2 \propto \tilde{Z}_0 \tilde{Z}_0^\dagger + \mathcal{O}(\tilde{\Delta}),
\]

(2.29)

where \( \tilde{\Delta} = e^{-(\bar{E}_1-\bar{E}_0)r} \sim e^{-(\bar{E}_1-\bar{E}_0)(t-r)} \). Considering the case \( \psi' = \psi \) and parametrizing with (2.21) we obtain

\[
\theta_*(\omega, \delta) = \theta_*(0,0) \frac{1 + 2\Delta \frac{x_{11}}{x_{00}} \tan \omega \cos \delta + \Delta^2 \frac{x_{11}^2}{x_{00}} \tan^2 \omega}{1 + \frac{\delta}{\omega} \tan^2 \omega + 2 \frac{\delta}{\omega} \tan \omega \cos(\arg(b) + \delta)},
\]

(2.30)

up to relative corrections of order \( \tilde{\Delta} \).

In the expression for two-point functions it was possible to find the location \( (\omega_*, \delta_*) \) of the global maximum analytically. However, this is not the case for three-point functions. Hence, we proceed by plotting the S/N ratio for the choice of parameters used for Fig. 2-2. The S/N landscape for three-point functions is shown in Fig. 2-3. From the figure it is clear that the S/N landscape is very similar to the one of two-point functions. Again, the symmetry about \( \delta = 0 \) is only due to the choice of parameter \( b \).

It turns out that for very specific cases \( (\chi_{11} = 0 \text{ and } \chi_{01} \neq 0) \) there can be two local maxima. However, for a physical system, \( \chi \) does not have this form, so there will always be a unique maximum. It is good to compare the location of the maximum S/N for two- and three-point functions. This is shown in Fig. 2-4. Note that the position of the maximum is generally different. This could be an issue since for the computation of \( R \) (defined in (1.19)) one needs both two- and three-point functions, so it would be desirable to optimize the S/N of both of them simultaneously. Fortunately, since \( \Delta \) is exponentially suppressed, then at large time separations the two- and three-point functions have a similar form, differing only up to an exponential factor. Hence, at late times, the location of the maximum should be fairly similar, which is indeed what is seen in Fig. 2-4.
2.3 Applications: Hadronic Three-Point Functions

We tested this technique with the computation of two observables: the axial charge of nucleons and $R$ (as defined in (1.19)) for the $\rho$ meson for an electromagnetic current with a nonzero momentum transfer. As mentioned before, we used Chroma to produce the gauge configurations and compute all quark propagators. For each test we used three different source-sink time separations: 15, 18, and 21, in order to understand excited state contamination. For the first two separations we used 200
gauge configurations and for the last one we used 500, since statistical noise increases with time separation. The lattices used were $16^3 \times 60$ with a bare anisotropy of $b_s/b_t = 2$, and a coupling of $\beta = 5.8$. The quarks masses used correspond to a pion mass of approximately 250 MeV$^1$.

The correlator matrices were constructed using 5 different interpolators. These consisted of a point interpolator, two Gaussian smearings of different widths, and the same Gaussians, but with a Laplacian operator applied to the smearing of one of the quarks. Following our arguments from above, we decided to find the variational and S/N optimized vectors only using the two-point functions.

To extract results from the plots of $R$ we fitted the data taking into account correlations between time slices. The ranges were chosen using an algorithm that determined the best plateau. The statistical uncertainty was found using statistical bootstrapping, and the systematic one was found using the standard deviation from small changes in the fitting range.

The correlator matrices are not perfectly hermitian or antihermitian. Since the S/N optimization only produces the same source and sink vectors when the matrix is hermitian, we enforce (anti)hermiticity by averaging over off-diagonal elements as necessary. This should not be a problem since in the limit of infinite statistics the matrices will be hermitian or antihermitian on average.

### 2.3.1 Nucleon Axial Charge

The axial charge of the nucleon can be calculated by considering the current with gamma structure $i\gamma_3\gamma_5$, and using a spin projector $\Gamma_3 = (\frac{1+i\gamma_4}{2})\gamma_3\gamma_5$, which corresponds to a polarized nucleon. The charge is found by computing $R$ as in Eq. (1.19) with momentum transfer $\vec{q} = 0$ and with $\vec{p}' = 0$. This observable provides a good test since it has been shown that the variational method can be used to achieve a more precise result [15]. Fig. 2-5 shows a comparison of the S/N optimization and

---

$^1$This value was obtained using a formula from Ref. [7] to convert the coupling to a lattice spacing, and then measuring the pion mass from an effective mass plot. The anisotropy of the lattice might have an effect on this estimate.
the variational method for a time separation of $t = 15$. From the figure it can be seen the values extracted from correlated fits are practically the same. This is also the case for the other time separations. A comparison of the values extracted from the fits of the variational and S/N optimized correlators, as well as the diagonal ones, is shown in Fig. 2-6.

![Figure 2-5: Comparison of the variational method, variational plus S/N optimization, and S/N optimization for the nucleon axial charge for a sink separation of $t = 15$. The S/N optimization offers no improvement over the variational method in this case.](image)

Figure 2-5: Comparison of the variational method, variational plus S/N optimization, and S/N optimization for the nucleon axial charge for a sink separation of $t = 15$. The S/N optimization offers no improvement over the variational method in this case.

![Figure 2-6: Comparison of the results obtained from the fits for the nucleon axial charge for different time separations. The S/N optimization offers no improvement over the variational method in this case.](image)

Figure 2-6: Comparison of the results obtained from the fits for the nucleon axial charge for different time separations. The S/N optimization offers no improvement over the variational method in this case.

We noticed that if we do not enforce (anti)hermiticity of the correlators, the S/N optimization gives a more precise result in some cases. The results in this scenario are shown in Fig. 2-7. It can be seen that for time separations $t = 15$ and $t = 18$ the S/N optimization offers slightly more precise results. However, it is also important to note that in these cases the extracted value is higher. This might be due to the loss
of positivity mentioned before. So although the result is more precise, it is likely less accurate and must be handled with care.

![Figure 2-7](image)

Figure 2-7: Same as Fig. 2-6, but without enforcing (anti)hermiticity. The S/N optimization for time separations $t = 15$ and $t = 18$ offers a slightly more precise result, but it is likely less accurate.

### 2.3.2 Rho Meson Form Factor

The other test for the effectiveness of the S/N optimization was to compute $R$ for the rho meson with a $\gamma_4$-current insertion with nonzero momentum transfer, which would be used in the computation of its electromagnetic form factor. We used a momentum transfer of $\vec{q} = (1, 0, 0)$, with the final rho at rest, $\vec{p}' = 0$. In this test the plots of $R$ are not symmetric since the source and sink momenta are different. Hence, the plateaus are not as pronounced as in the previous case. This caused our algorithm to fail in finding a plateau in some cases, so the fitting ranges had to be chosen manually. Fig. 2-8 shows a comparison of the S/N optimization and the variational method for a time separation of $t = 15$. We again compile the results from all correlators, and plot them in Fig. 2-9. In this case, there is a very slight improvement in the precision of the $t = 15$ result. However, it is virtually the same for the other two time separations.
Figure 2-8: Comparison of the variational method, variational plus S/N optimization, and S/N optimization for the rho form factor for a sink separation of $t = 15$. The S/N optimization offers a very slight improvement.

Figure 2-9: Comparison of the results obtained from the fits for the rho form factor for different sink separation times. The S/N optimization offers only a very slight improvement for $t = 15$.

We again found that without enforcing hermiticity, the noise is reduced further. The effect is quite significant for each individual data point, as can be see in Fig. 2-10. However, the plateaus are even more subtle. Therefore, the systematic uncertainty becomes larger, inhibiting the reduction of the statistical error. This can be seen in the comparison in Fig. 2-11. Furthermore, we again see that the loss of positivity has a nonnegligible effect, as the results for the S/N optimized correlators are lower than the other ones.
Figure 2-10: Same as Fig. 2-8, but without enforcing hermiticity. The S/N optimization significantly reduces the noise of individual data points, but the plateaus become less pronounced.

Figure 2-11: Same as Fig. 2-9, but without enforcing hermiticity. The S/N optimization offers practically no improvement overall, since the more subtle plateaus inhibit the improvement in the statistical noise.

2.4 Discussion and Conclusions

From these two tests we saw that there was no case where the S/N optimization offered a clear advantage over the variational method. When enforcing hermiticity, the results from the S/N optimization are virtually the same as those obtained with the variational method. When hermiticity is not enforced, the noise is reduced, but at the expense of less pronounced plateaus and likely less accurate results due to loss of positivity.

In the original paper by Detmold and Endres, the S/N optimization provided a clear improvement in only one of the cases. So it was concluded that the method may have to be compared to the variational method on a case-by-case basis to determine
its benefit. The conclusion here seems to be the same. Although we could not show its effectiveness, it may be the case that for other observables or in other settings it can be superior to the variational method.

In this study we used relatively basic quenched configurations. Besides, the S/N optimization did not significantly reduce the noise in effective mass plots. In contrast, Detmold and Endres used data from unquenched configurations and using improved actions. They obtained much more noticeable improvements in the noise of two-point function. Hence, this procedure will also depend on the configurations used. It is possible that for more carefully constructed configurations, the noise in three-point functions can see much better improvements from the S/N optimization than those shown here. Further studies need to be done to test whether this is the case.
Chapter 3

Gauge Field Data Compression

As we discussed in Sec. 1.3, the computation of observables in Lattice QCD requires a large ensemble of gauge configurations. The tremendous advances in computing power seen in the last decades have allowed for the use of much larger lattices and ensembles. The increase in lattice sizes makes them more difficult to store and transfer. For instance, a current problem is the narrow bandwidth available for the transmission of gauge configurations to GPUs. It is thus desirable to use a data compression method to store gauge configurations more efficiently. However, due to the stochastic nature of the configurations and issues such as gauge symmetry, standard compression algorithms are of no use.

There are two trivial methods to compress $SU(3)$ fields. The first one is by exploiting the fact that $SU(3)$ matrices have the form

$$M = \begin{pmatrix} u & \ast \times v^* \\
                          v & \end{pmatrix}.$$  \hspace{1cm} (3.1)

Therefore, one can store only the first two rows, and the last one can be reconstructed when needed. Thus, this requires $12/18$ of the original space. A similar strategy can also be applied to other $SU(N)$ matrices. The other method, which works for general $SU(N)$ matrices, uses the Lie algebra structure of the group. An arbitrary matrix
\( M \in SU(N) \) can be written as

\[
M = \exp(ic_n T_n),
\]

(3.2)

where \( T_n \) are the generators of the \( su(N) \) algebra, and \( c_n \) are some coefficients (we use the Einstein summation convention). By storing only the coefficients \( c_n \) it requires 8/18 of the original space when \( N = 3 \).

There is yet another method to compress gauge fields, which relies in gauge invariance. One can gauge fix with a temporal gauge, so that all temporal links are set to the identity, reducing the storage needed to 3/4. Since this method can be applied concurrently with ones that compress individual \( SU(N) \) matrices, we will focus on the other ones for the rest of this chapter.

In this chapter we explore a compression method that could reduce the space needed even further than the Lie algebra compression. The idea is to discretize the \( SU(N) \) manifold very finely and store only a number corresponding to the matrix closest to the given matrix. This approach is not new. The first study of discretized subsets of gauge groups was done by Petcher and Weingarten for the case of \( SU(2) \) [16]. They found that the physics is well approximated when only considering its largest finite subgroup, the icosahedral group \( I \) of order 120 (we refer to it as \( G_{120} \)). Lisboa and Michael later studied the case of \( SU(3) \) [17], which is relevant for QCD. They found that its largest finite subgroup \( G_{1080} \) of order 1080 failed to approximate the physics at the scale of interest. They instead proposed to use a finite set \( M_{1080} \), with 38,880 elements, composed of the midpoints of the nearest neighbors in \( G_{1080} \). It was shown that \( M_{1080} \) served as a good approximation at the scale of interest. Although the method by Lisboa and Michael provides a good approximation, the compression is far from being lossless. We explore a method that could offer machine precision at a fraction of the storage needed by other approaches. Note that “precision” usually refers to the number of significant digits, but since the matrix elements in this case are generally of order 0.1, then we use “single precision” to denote a difference of less than \( 10^{-7} \) and “double precision” for a difference of less than \( 10^{-16} \).
3.1 Compression Method Description

The compression method we explore was proposed by William Detmold and Michael Endres [18] following the ideas of the group discretizations mentioned above. The idea is to first use a finite subset $\Sigma$ of elements of the group that are distributed along the manifold in a roughly uniform manner, such as $G_{120}$ for $SU(2)$ or $M_{1080}$ for $SU(3)$. Then, around each element of $\Sigma$ one constructs a very fine lattice of random elements. To do this, one picks $n$ random matrices $R_i$ that are close to the identity. One then considers the products

$$R_1^{b_1} R_2^{b_2} \cdots R_n^{b_n},$$  \hspace{1cm} (3.3)

where the exponents $b_i$ are either zero or one. These products form a fine irregular lattice of matrices near the identity. One can then translate this lattice to each element $S \in \Sigma$ by multiplying by $S$. The spread of the random matrices is controlled so that when the lattices around all elements of $\Sigma$ are considered, the whole manifold is covered, yet there is not too much overlap between the lattices of nearest neighbors in $\Sigma$. Thus, given a matrix $M$, it can be approximated by first finding the closest element $S$ of $\Sigma$, and then the closest element in the corresponding fine lattice,

$$M \approx SR_1^{b_1} R_2^{b_2} \cdots R_n^{b_n}. \hspace{1cm} (3.4)$$

Due to the random nature of the matrices $R_i$ it is (almost always) the case that given a plane passing though the origin (the identity) there will be more matrices on one side than in the other. This results in the distribution of products being skewed to one of the sides. This can also be noticed when orthants are used, since opposite orthants tend to have a very different number of elements. We found that a good solution for this is to remove one of the matrices $R_i$ and instead use the corresponding bit to denote Hermitian conjugation, which effectively reflects the matrices about the origin. This makes the fine lattice much more uniform overall. It is again translated
to all elements of $\Sigma$ so that Eq. (3.4) is modified to become

$$M \approx \begin{cases} S R_i^b R_2^b \cdots R_{n-1}^{b_{n-1}} & \text{if } b_n = 0, \\ S(R_i^b R_2^b \cdots R_{n-1}^{b_{n-1}}) & \text{if } b_n = 1. \end{cases} \quad (3.5)$$

Suppose that $n'$ bits are required to enumerate the elements of $\Sigma$. Since each $b_i$ corresponds to a bit, then $n + n'$ bits are required to specify the matrix $M$.

It is essential to define a metric on the $SU(N)$ manifold, in order to find the closest element in the lattice to a given matrix $M$. The metric is defined so that the distance between two matrices $M_1$ and $M_2$ is given by

$$\text{dist}(M_1, M_2) = \sqrt{\frac{1}{2} \sum_i \arg(\lambda_i(M_1 M_2)^2)}, \quad (3.6)$$

where $\lambda_i(M)$ are the eigenvalues of $M$.

If $\Sigma$ a is large enough set, then the space around each of its element is approximately flat. The distance (3.6) corresponds to the norm of the vector $\vec{c}$ (as in (3.2)) of coefficients of the Lie algebra generators (if $|\vec{c}| \leq 2\pi$). For nearby matrices, this distance is roughly the same as the maximum difference between their matrix elements. This is particularly clear for $SU(2)$ since we have the identity

$$e^{i\theta(\hat{n} \cdot \vec{c})} = I \cos \theta + i(\hat{n} \cdot \vec{c}) \sin \theta, \quad (3.7)$$

where $\hat{n}$ is a unit vector, and $\theta$ corresponds to the distance given by (3.6). Thus, for small distances the matrix elements differ by at most $\sin \theta \approx \theta$. For $SU(3)$ we note that seven of the Gell-Mann matrices have the form of $SU(2)$ matrices with some extra row and column inserted symmetrically. Thus, the above formula gives a similar result for each one. The last matrix is diagonal, so it can easily be checked to behave similarly. Furthermore, since the vector $\vec{c}$ generally points in a random direction, the differences between matrix elements will generally be smaller than the distance between the matrices. Therefore, the distance given by (3.6) roughly corresponds to the maximum difference between the matrix elements. Thus, we can use the distance
between the given matrix $M$ and the one constructed with (3.5) to quantify the precision of the approximation.

### 3.2 Theoretical performance

We now estimate the best possible compression that can be achieved with this method. The $SU(N)$ group manifold has dimension $N^2 - 1$. Let us assume that the manifold is covered by a fine enough lattice so that it can locally be approximated by flat space. Let us also assume that the lattice used to cover the manifold is locally (hyper)cubic. We first consider the case for $SU(2)$. The group manifold is a 3-sphere. The above metric corresponds to the sphere having unit radius, so it has a cubic hyperarea of $2\pi^2$. If we want to achieve machine precision ($10^{-10}$) then this must be the farthest possible distance between a given point and a node of the lattice. This would correspond to the center of a cubic element, so the lattice constant must be $a = \frac{2}{\sqrt{3}} \cdot 10^{-7}$. Thus, the number of bits required to enumerate (and hence store) matrices in the lattice is

$$N_{\text{bits } SU(2)} = \log_2 \left( \frac{2\pi^2}{8 \left( \frac{2}{\sqrt{3}} \cdot 10^{-7} \right)^3} \right) = 70.441... \approx 71, \quad (3.8)$$

where the factor of 8 comes from the fact that each vertex is shared by 8 cubic cells.

The $SU(3)$ manifold is non-trivial. We will assume that it has the same hyperarea as a unit 8-sphere to compute an analogous estimate, although it is unclear whether this is a good approximation. With similar arguments as above we obtain that the number of bits required to store $SU(3)$ matrices with machine precision is

$$N_{\text{bits } SU(3)} = \log_2 \left( \frac{\frac{\pi^4}{3}}{256 \left( \frac{2}{\sqrt{8}} \cdot 10^{-7} \right)^8} \right) = 187.049... \approx 187. \quad (3.9)$$

The bounds given by Eqs. (3.8) and (3.9) could be slightly lowered if one uses a face-centered cubic lattice for $SU(2)$ and an $E_8$ lattice for $SU(3)$, which would provide the most efficient packing for $\mathbb{R}^3$ and $\mathbb{R}^8$ respectively [19]. Nevertheless, these
estimated limits provide us with a value that can be compared with our results below.

We can also estimate the dependence of the precision on the number of bits. To decrease the distances in a lattice by a factor of 10, the number of nodes must be increased by $10^d$, where $d$ is the dimension. This corresponds to adding a number of bits $\delta n$ given by

$$\delta n = d \cdot \log_2(10).$$

(3.10)

For $SU(2)$ this corresponds to an increase of $\delta n = 9.966...$ bits, and for $SU(3)$ it corresponds to $\delta n = 26.575...$ bits. Again, this might differ from the observed value, but it gives us an idea of the expected behavior.

### 3.3 Implementation

One might be inclined to pick the matrices $R_i$ in a systematic way, such as using the vertices of a regular polytope. However, the symmetry from these structures results in many cancellations between the matrices, which makes the products very poorly distributed. Hence, it is essential to use an asymmetric choice of matrices $R_i$. Thus, as mentioned above, these matrices are chosen randomly. Although there are several possible ways of doing this, we found that the procedure giving the most consistent results is to pick the matrices from a uniform sphere of radius $\epsilon$ centered at the origin (the identity). Then $\epsilon$ can be adjusted to regulate the spread of the products (3.3).

In order to choose the parameter $\epsilon$ one can start by plotting the distance of random $SU(N)$ matrices to the closest element in the set $\Sigma$. Fig. 3-1 shows an example for $SU(2)$ with the set $\Sigma$ chosen to be $G120$.

One can then generate various sets of matrices $R_i$ with different parameter $\epsilon$ and plot the distance between their products to the origin. These distributions are roughly approximated by a Gaussian, but they have positive skewness. Some examples are shown Fig. 3-2 for the case of $SU(2)$. Comparing Figs. 3-1 and 3-2 one sees that the shapes are very different, which means that the distribution of the matrices will not be ideal. In particular, many of the matrices farther from the identity will remain unused. Nevertheless, one can pick the one that most closely resembles the shape and
peak location of the distribution in Fig. 3-1, which if done correctly, should result in a lower average distance between a given random matrix and one approximated using Eq. (3.5).

The next important step is to devise a method to find the closest element obtained from Eq. (3.5) to a given matrix. One can first perform a direct search to find the closest \( S \in \Sigma \) since this set is not very large. To search the corresponding fine lattice of \( S \) we tried two methods: multilateration and a tree search. For the multilateration we stored the distances from all matrices to \( d + 1 \) fixed matrices. The distances were stored in ascending order to speed up the search. Thus, given a matrix \( M \), one can compute the distances to the fixed matrices, take some elements in a small range for each fixed points, and then take the intersection of these sets. In this way one only
has to search a small portion of the lattice, which can be done very quickly. The problem we encountered is that taking the intersections of these sets is extremely slow, so this method was discarded. Instead, we used a tree search. We partitioned the space of matrices and stored the matrices in each section. Thus, given a matrix $M$, we first determine which section it is in, and then perform a direct search for all the matrices in that section. It could be extended even further by subdividing each section. However, since the closest matrix may lie on a neighboring section, then one must be careful to not divide the space too finely, or take precautions such as constructing overlapping sections.

3.4 Application: SU(2) Field

We focused on the case of $SU(2)$ matrices. We used the icosahedral group $G_{120}$ as our set $\Sigma$, and we considered the case of 16, 18, 20, 22, and 24 bits for the fine sublattice. However, due to the exponential increase in time it became unmanageable to study finer lattices. For 16, 18, and 20 bits we partitioned the space around each element of $\Sigma$ into octants. For 22 and 24 bits we divided the space for the fine search according to the distance $r$ between the given $M$ and the closest $S \in \Sigma$. For $r < 0.1$ the matrices are grouped together; for $0.1 < r < 0.2$ the matrices are grouped according to the faces of a cube; for $0.2 < r < 0.3$ according to the faces of a dodecahedron; and for $0.3 < r$ according to the faces of an icosahedron. This partition was chosen so that for larger radii there were more subdivisions since there are more matrices.

For each number of bits we picked 5 sets of matrices. The choice had to be done manually, since by the random nature of the procedure some sets are more suitable than others, even when using the same parameter $\epsilon$. We then generated 5 ensembles of 1000 random $SU(2)$ matrices and computed the distance to the closest element in the lattice. The merged distributions are shown in Fig. 3-3.

From Fig. 3-3 is is clear that the average distance to the closest element in the lattice is reduced as the number of bits is increased. Furthermore, the tail of the distribution is significantly improved as well. This is important since for a lossless
Figure 3-3: Distance from random matrices to the closest matrix in the lattice. Increasing the number of bits reduces the average distance, as well as improving the behavior of the tail.

In compression it is necessary to not have outliers. We can quantify the quality of the tails by computing the kurtosis. More specifically, we use the excess kurtosis, which is defined as kurtosis minus three. This is shown in Table 3.1. The table also shows the average and maximum distance to the closest element in the lattice, as well as the time taken to find it.

Table 3.1: Results from using 5 sets of matrices $R_i$ and 1000 random $SU(2)$ matrices. The increase in kurtosis and decrease in elapsed time at 22 bits is due to the switch to a different partitioning.

<table>
<thead>
<tr>
<th>Bits</th>
<th>Average [rad]</th>
<th>Maximum [rad]</th>
<th>Excess kurtosis</th>
<th>Average [s]</th>
<th>Maximum [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>$(96.4 \pm 3.6) \cdot 10^{-4}$</td>
<td>$(38.2 \pm 5.3) \cdot 10^{-3}$</td>
<td>$3.87 \pm 2.04$</td>
<td>$0.62 \pm 0.05$</td>
<td>$1.55 \pm 0.37$</td>
</tr>
<tr>
<td>18</td>
<td>$(61.2 \pm 5.0) \cdot 10^{-4}$</td>
<td>$(22.4 \pm 9.4) \cdot 10^{-3}$</td>
<td>$2.89 \pm 3.88$</td>
<td>$2.71 \pm 0.22$</td>
<td>$6.03 \pm 0.92$</td>
</tr>
<tr>
<td>20</td>
<td>$(37.4 \pm 0.6) \cdot 10^{-4}$</td>
<td>$(13.4 \pm 1.4) \cdot 10^{-3}$</td>
<td>$2.00 \pm 0.65$</td>
<td>$11.98 \pm 0.99$</td>
<td>$21.70 \pm 4.67$</td>
</tr>
<tr>
<td>22</td>
<td>$(24.7 \pm 0.8) \cdot 10^{-4}$</td>
<td>$(10.9 \pm 2.9) \cdot 10^{-3}$</td>
<td>$5.45 \pm 5.74$</td>
<td>$11.67 \pm 1.07$</td>
<td>$27.78 \pm 4.29$</td>
</tr>
<tr>
<td>24</td>
<td>$(14.2 \pm 0.3) \cdot 10^{-4}$</td>
<td>$(4.5 \pm 1.0) \cdot 10^{-3}$</td>
<td>$1.38 \pm 1.70$</td>
<td>$52.58 \pm 5.31$</td>
<td>$109.85 \pm 24.97$</td>
</tr>
</tbody>
</table>

From the table it is clear that there is indeed a downward trend for the excess kurtosis, but it is unclear whether it will approach zero (the kurtosis of a Gaussian). Nevertheless, this shows that the tails are reasonably well behaved as the number of bits is increased. Note that the increase in kurtosis and decrease in time at 22 bits is due to the switch to a different partitioning scheme, as mentioned earlier. This confirms that one has to be careful not to divide the space too finely, or to take precautions before doing so.
We can now compare the behavior of the precision as a function of the number of bits. We do so by plotting the average and maximum distances and fitting with an exponential function. This is shown in Fig. 3-4.

![Figure 3-4: Average and maximum distance between random SU(2) matrices and the closest element in the lattice as a function of bits used for the fine lattice. The lines correspond to the best exponential fits.](image)

From Fig. 3-4 one can see that there is a very nice exponential behavior for the average and maximum distances. From the fit we obtain that the average is decreased by an order of magnitude by adding $\delta n_{\text{avg}} = 9.65 \pm 0.04$ bits and the maximum is decreased by an order of magnitude by adding $\delta n_{\text{max}} = 8.86 \pm 0.22$ bits. This is fairly close to what was predicted with Eq. 3.10. Presumably, it is slightly lower because by increasing the number of bits the lattice also becomes more uniform. Using the slope of the fit from the averages (since it is more accurate), we calculate the number of bits required to make the maximum distance be $10^{-7}$. We obtained that around 70 bits (for the fine lattice) are required. Since the elements of $G_{120}$ can be enumerated with 7 bits, then 77 bits in are required in total to describe the matrix. This result is shown in Table 3.2, where it is also compared with other compression methods and with the estimated theoretical limit.

From Table 3.2 we see that this method provides a compression that is quite close to the theoretically best compression that we determined earlier. These estimates, especially the one for double precision, must be taken with some degree of skepticism since we are extending the results from Fig. 3-4 very far into the right. It could be the case that the slope becomes closer to the one predicted with Eqn. (3.10) since the
Table 3.2: Comparison of the estimated number of bits required to describe a matrix with this method and other methods, as well as the best theoretical compression.

<table>
<thead>
<tr>
<th>Precision</th>
<th>Number of bits required</th>
<th>Full matrix</th>
<th>First row</th>
<th>Lie algebra</th>
<th>Discretization</th>
<th>Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Half</td>
<td></td>
<td>128</td>
<td>64</td>
<td>48</td>
<td>38</td>
<td>31</td>
</tr>
<tr>
<td>Single</td>
<td></td>
<td>256</td>
<td>128</td>
<td>96</td>
<td>77</td>
<td>71</td>
</tr>
<tr>
<td>Double</td>
<td></td>
<td>512</td>
<td>256</td>
<td>192</td>
<td>164</td>
<td>160</td>
</tr>
</tbody>
</table>

lattice stops becoming more uniform at some point. Nevertheless, since the lattice is substantially nonuniform for the number of bits we explored, we assume that the point where it stops becoming more uniform is even farther into the right.

Note that we ignored the fact that the kurtosis has a downward trend, so the outliers become less pronounced and the maximum distance decreases even faster than the average. Furthermore, if one accepts having a few outliers, then it would be possible to use only 64 bits to represent the matrix at close to single precision, which would be very convenient. Thus, this compression reduces the size to \( \frac{2}{3} \) compared to the next best one, while only being an order of magnitude off from single precision.

### 3.5 Application: SU(3) Field

We did not perform such a thorough analysis for \( SU(3) \) matrices, but we estimated how they should behave. We used the midpoints \( M_{1080} \) as our set \( \Sigma \). With 16 bits for the fine lattice we found that both the average and maximum distance were roughly a factor of 25 larger than what we found for \( SU(2) \). Assuming value for \( \delta n \) slightly lower than what is found with Eq. (3.10) (to account for the increase in uniformity), we obtain that 170 bits are required for the fine lattice to achieve single precision. Since 16 bits are required to enumerate the elements of \( M_{1080} \), then 186 bits are requires in total. We show a comparison in Table 3.3.

We again see that this method provides a compression that is relatively close to the best compression possible. However, the size is not reduced as much as in the case of \( SU(2) \).
### Table 3.3: Same as Table 3.2, but for SU(3).

<table>
<thead>
<tr>
<th>Precision</th>
<th>Number of bits required</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full matrix</td>
</tr>
<tr>
<td>Half</td>
<td>288</td>
</tr>
<tr>
<td>Single</td>
<td>576</td>
</tr>
<tr>
<td>Double</td>
<td>1152</td>
</tr>
</tbody>
</table>

### 3.6 Discussion and Conclusions

We found that it is indeed possible to achieve a better compression by discretizing the group manifold. The decompression is clearly very easy since it only involves multiplication of matrices, which can be done very fast. The remaining question is whether it is worthwhile or even feasible to compress the data. The size reduction is not as significant as one might initially expect, but we showed that this method gets close to the best compression possible by this type of approaches.

There are two main issues that this method faces. One of them is that the best compression seems to get closer to the Lie algebra one as the precision is increased. This makes it less ideal for high precision studies, which presumably would be the most benefited by data compression. The other, and main problem is that the compression is extremely costly. From Table 3.1 one sees that compressing a single matrix for a (fine) lattice constructed with 24 bits takes on average 53 seconds. A gauge configuration would have hundreds of thousands or even millions of matrices, making the total time needed impractical. Furthermore, to achieve single precision a much higher number of bits is required. Therefore, it is currently completely impractical unless a much more efficient way of finding the closest matrix is devised.

There are other possible improvements that can be made when storing entire gauge fields or ensembles of gauge configurations. As the continuum limit is approached, the gauge fields becomes smooth, and thus, matrices from nearby links will be similar. Hence, it might be possible to store an average matrix for a small region, and specify small adjustments for individual links with only a few bits. This would remove the need to have a machine precision coverage of the entire manifold, so it would be much
more practical. Moreover, what is important is only for calculation of observable quantities to be correct. Thus, it may be that for a given ensemble of configurations, each configuration need not be stored at machine precision to achieve an equally good result. Investigations of such strategies are left to future work.
Chapter 4

Conclusions

In Chapter 2, we explored a method to reduce the noise in three-point correlators. It was an extension from a method proposed by Detmold and Endred to improve the S/N ratio of two-point functions [13]. We tested the technique in two cases where it could potentially be useful: the measurement of the nucleonic axial charge, and in a matrix element computation related to the electromagnetic form factor of the rho meson. For the case of the nucleon we did not see any improvement over the variational method. In the case of the rho, there was a very slight improvement. We saw a more significant improvement in both cases when (anti)hermiticity is not enforced. However, this result is likely less accurate due to loss of positivity. We concluded that a case-by-case study would be needed to determine the effectiveness of the S/N optimization over the variational method. We also concluded that if more sophisticated configurations are used (i.e. unquenched and with improved actions), then the improvement in the S/N ratio will possibly be more noticeable.

In Chapter 3, we explored a compression method that could offer a better compression ratio than storing Lie algebra coefficients. It consisted in generating a set $\Sigma$ of matrices distributed roughly uniformly along the manifold, and constructing fine lattices around each of these elements. We showed that this compression can indeed provide better performance for $SU(2)$ and $SU(3)$ than the Lie algebra compression, but the improvement is not very significant. Nevertheless, it is fairly close to the best compression one could hope to achieve with this type of method. Another problem
was that the compression is very costly, which makes it currently impractical if one desires to achieve machine precision. We concluded that it must be studied whether additional information, such as correlations in the lattice, can be used to improve the compression and make the method more practical.
Bibliography


59


