

2017

Disconnected Diagrams in Lattice Qcd

Arjun Singh Gambhir

College of William and Mary, asgambhir@email.wm.edu

Follow this and additional works at: <https://scholarworks.wm.edu/etd>



Part of the [Atomic, Molecular and Optical Physics Commons](#)

Recommended Citation

Gambhir, Arjun Singh, "Disconnected Diagrams in Lattice Qcd" (2017). *Dissertations, Theses, and Masters Projects*. Paper 1516639772.

<http://dx.doi.org/doi:10.21220/S2108D>

This Dissertation is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Dissertations, Theses, and Masters Projects by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.

Disconnected Diagrams in Lattice QCD

Arjun Singh Gambhir

Fairfax, Virginia

Master of Science, College of William & Mary, 2013
Bachelor of Science, Virginia Polytechnic Institute and State University, 2011

A Dissertation presented to the Graduate Faculty
of The College of William & Mary in Candidacy for the Degree of
Doctor of Philosophy

Department of Physics

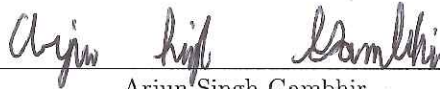
College of William & Mary
August 2017

©2017
Arjun Singh Gambhir
All rights reserved.

APPROVAL PAGE

This Dissertation is submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy



Arjun Singh Gambhir

Approved by the Committee, June, 2017



Committee Chair

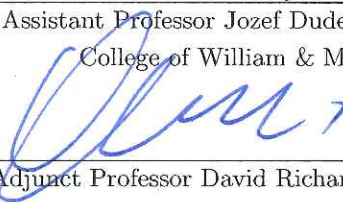
Professor Konstantinos Orginos, Physics
College of William & Mary



Professor Carl Carlson, Physics
College of William & Mary



Assistant Professor Jozef Dudek, Physics
College of William & Mary



Adjunct Professor David Richards, Physics
College of William & Mary



Professor Andreas Stathopoulos, Computer Science
College of William & Mary

ABSTRACT

In this work, we present state-of-the-art numerical methods and their applications for computing a particular class of observables using lattice quantum chromodynamics (Lattice QCD), a discretized version of the fundamental theory of quarks and gluons. These observables require calculating so called “disconnected diagrams” and are important for understanding many aspects of hadron structure, such as the strange content of the proton. We begin by introducing the reader to the key concepts of Lattice QCD and rigorously define the meaning of disconnected diagrams through an example of the Wick contractions of the nucleon. Subsequently, the calculation of observables requiring disconnected diagrams is posed as the computationally challenging problem of finding the trace of the inverse of an incredibly large, sparse matrix. This is followed by a brief primer of numerical sparse matrix techniques that overviews broadly used methods in Lattice QCD and builds the background for the novel algorithm presented in this work. We then introduce singular value deflation as a method to improve convergence of trace estimation and analyze its effects on matrices from a variety of fields, including chemical transport modeling, magnetohydrodynamics, and QCD. Finally, we apply this method to compute observables such as the strange axial charge of the proton and strange sigma terms in light nuclei.

The work in this thesis is innovative for four reasons. First, we analyze the effects of deflation with a model that makes qualitative predictions about its effectiveness, taking only the singular value spectrum as input, and compare deflated variance with different types of trace estimator noise. Second, the synergy between probing methods and deflation is investigated both experimentally and theoretically. Third, we use the synergistic combination of deflation and a graph coloring algorithm known as hierarchical probing to conduct a lattice calculation of light disconnected matrix elements of the nucleon at two different values of the lattice spacing. Finally, we employ these algorithms to do a high-precision study of strange sigma terms in light nuclei; to our knowledge this is the first calculation of its kind from Lattice QCD.

TABLE OF CONTENTS

Acknowledgments	v
Dedication	vi
List of Tables	vii
List of Figures	viii
CHAPTER	
1 Motivation: Hadron Structure	2
1.1 Proton spin crisis	5
1.2 Dark matter cross sections	8
1.3 Proton radius puzzle	10
1.3.1 Strange radius	12
2 Introduction to Lattice QCD	14
2.1 Continuum QCD	15
2.1.1 Fields in the QCD action	15
2.1.2 The fermionic QCD action	16
2.1.3 Verifying Gauge invariance	17
2.1.4 The gluonic QCD action	19
2.1.5 Asymptotic freedom	21
2.2 The Wilson pure gauge action	22
2.2.1 Gauge fields on the lattice as link variables	22
2.2.2 Link variables in the continuum	23
2.2.3 Gauge-invariant objects on the lattice	24

2.2.4	The Wilson gauge action	25
2.3	Fermions on the lattice	27
2.3.1	Naive lattice fermions	27
2.3.2	The Dirac operator and the doubling problem	29
2.3.3	Wilson fermions	32
2.4	Symanzik improvement	33
2.4.1	Toy model	34
2.4.2	Implementation to Lattice QCD	36
3	Computing Matrix Elements in Lattice QCD	39
3.1	Two-point functions	40
3.1.1	Nucleon interpolating field	40
3.1.2	Contractions	42
3.1.3	Quark sources and smearing	43
3.2	Three-point functions	44
3.2.1	Up quark contractions	45
3.2.2	Down quark contractions	47
3.2.3	Strange quark contractions	49
3.2.4	Difference between connected and disconnected diagrams	50
3.3	Reducing excited state contamination	51
3.3.1	Plateau method	52
3.3.2	Two-state model	53
3.3.3	Exponential Fit	54
3.3.4	Summation method	55
4	Introduction to Numerical Linear Algebra for Sparse Matrices	57
4.1	Preamble	58

4.1.1	Direct vs iterative methods	58
4.1.2	Condition number and numerical stability	61
4.2	Linear solvers	62
4.2.1	Conjugate Gradient	63
4.2.2	Bi-conjugate gradient stabilized method	67
4.2.3	Preconditioners and multigrid	69
4.3	Trace estimation techniques	71
4.3.1	Hutchinson trace	72
4.3.2	Probing	74
4.3.3	Hierarchical probing	75
5	SVD Deflation as a Method of Variance Reduction For Estimating the Trace of a Matrix Function	79
5.1	SVD deflation and its impact on variance	80
5.2	The effect of the singular spectrum	84
5.3	Experiments on general matrices	87
5.3.1	Experiments with \mathbb{Z}_4 noise	87
5.3.2	Experiments with Gaussian vectors	89
5.4	Application to Lattice QCD	93
5.4.1	How to obtain the deflation space	93
5.4.2	SVD deflation and HP trace algorithm	95
5.5	Numerical experiments in Lattice QCD	96
5.5.1	Synergy between deflation and hierarchical probing	97
5.5.2	Varying the SVD deflation space	101
5.5.3	Wallclock timings and efficiency	102
6	Selected Results	106
6.1	Varying HP vectors	107

6.1.1	Ensemble and correlator overview	107
6.1.2	Axial charge	109
6.1.3	Scalar charge	114
6.1.4	Tensor charge	117
6.1.5	Vector charge	120
6.2	Disconnected nucleon matrix elements	124
6.2.1	Preliminaries	124
6.2.2	Axial charge	126
6.2.3	Scalar charge	129
6.2.4	Renormalized charges	133
6.3	High-precision calculation of strange sigma terms in light nuclei	135
6.3.1	Lattice parameters and setup	135
6.3.2	Proton-proton	136
6.3.3	Deuteron	140
6.3.4	^3He	142
6.3.5	Final results	146
7	Summary	148
APPENDIX A		
	Euclidean Gamma matrices	151
APPENDIX B		
	Grassmann numbers	154
APPENDIX C		
	A closer look at conjugate gradient	158
	Bibliography	161

ACKNOWLEDGMENTS

I owe a great deal to many people for aiding with research and making my time in graduate school some of the best years of my life. My family and friends have provided immeasurable support during my time at William and Mary. I want to thank Boram Yoon, Rajan Gupta, and the NME collaboration for sharing their two-point functions and “naive” disconnected quark loops for use as a comparison with our methods. I thank David Richards for discussions, advice, and teaching a wonderful class that has clarified my understanding of Lattice QCD. During his brief time at Jefferson Lab and thereafter, Sergey Syritsyn has always been willing to answer questions, discuss a new idea, or help optimize our codes, for that I am grateful. I cannot say enough about Bálint Joó, whose invaluable contributions extend far beyond helping me. I am confident that without him, Jefferson Lab’s lattice effort would be half as good at best.

I want to thank Andreas Stathopoulos, who has always been willing to discuss, be it day or night, in person or remotely. He has shown me that research is not a career, but an incredible hobby. I am most grateful to my mentor and teacher, Kostas Orginos. I want to thank him for giving me the passion to explore new ideas and for being an amazing lifeline of knowledge and support when deadlines were fast approaching. However, I want to thank him even more for the many times in which he suggested that I explore independently and figure things out for myself. Through his mentorship I have gained the confidence and resourcefulness to stand on my own. Finally, I would like to acknowledge Andreas and Kostas not individually, but as a team: from your example I have learned that the best aspect of research is not the work we do, but the friendships we make along the way...

I would like to dedicate this work to my family to honor their unending love and support. To my parents, Ajit and Ravinder Gambhir, thank you for your sacrifices in ensuring educational opportunities for me that you did not have. To my brother, Harman Ahuja, your exemplary work in school and toward your career has taught me much. Thank you.

LIST OF TABLES

6.1	Ensemble Parameters	125
6.2	Quark Loop Parameters	125
6.3	Nucleon Correlator Parameters	125
6.4	Proton-Proton/Proton Exponential Fit Result	137
6.5	Proton-Proton/Proton Exponential Fit χ^2 /d. o. f.	137
6.6	Proton-Proton/Proton Summation Fit Result	137
6.7	Proton-Proton/Proton Summation Fit χ^2 /d. o. f.	138
6.8	Proton-Proton/Proton Two-State Covariant Fit Result	138
6.9	Proton-Proton/Proton Two-State Covariant Fit χ^2 /d. o. f.	139
6.10	Deuteron/Proton Exponential Fit Result	140
6.11	Deuteron/Proton Exponential Fit χ^2 /d. o. f.	140
6.12	Deuteron/Proton Summation Fit Result	140
6.13	Deuteron/Proton Summation Fit χ^2 /d. o. f.	141
6.14	Deuteron/Proton Two-State Covariant Fit Result	141
6.15	Deuteron/Proton Two-State Covariant Fit χ^2 /d. o. f.	142
6.16	^3He /Proton Exponential Fit Result	143
6.17	^3He /Proton Exponential Fit χ^2 /d. o. f.	143
6.18	^3He /Proton Summation Fit Result	144
6.19	^3He /Proton Summation Fit χ^2 /d. o. f.	144
6.20	^3He /Proton Two-State Covariant Fit Result	145
6.21	^3He /Proton Two-State Covariant Fit χ^2 /d. o. f.	145

LIST OF FIGURES

2.1	Feynman diagrams representing the gluon self interaction terms found in the QCD action. Unlike photons in QED, gluons carry color charge, which engenders the non-linear nature of the strong force.	20
2.2	A pictorial representation of $Q_{\mu\nu}$ in the $\mu - \nu$ plane.	38
3.1	Here we schematically show the contractions of the nucleon two-point correlator. The top fermion line is the down quark, the bottom two are up quarks.	43
3.2	The first four terms of the up quark contractions in (3.16) are represented pictorially above. These diagrams only illustrate the Wick contractions, as all possible gluons connect the fermion lines. The top fermion is the down quark and the bottom two are up quarks in the proton. These diagrams are typically classified as the connected pieces.	46
3.3	The last two terms of (3.16) are shown schematically. The loop is comprised of an up quark propagating from around the same timeslice. These are categorized as the disconnected contributions of the up quark.	47
3.4	The first two terms in (3.17) are shown. Since the proton has only one down quark, only two different contractions are present in the connected case.	48
3.5	Here we show the last two terms of (3.17). The loop is made of a propagating down quark. Due to isospin symmetry, usually only one computation is done for Figures 3.3 and 3.5.	48
3.6	The only types of contributions from strange quark contractions are disconnected.	49
5.1	On the left is a logarithmic spectrum: $\sigma_{N-i+1} = 1 + 2 \cdot \log(i)$. On the right is a square root spectrum: $\sigma_{N-i+1} = \sqrt{i}$. The dotted red line in both plots is a constant line at $y = 1$. Points below this line signify an improvement in variance with deflation. Points above the line denote a deflated operator with a higher Frobenius norm than the original matrix, a case in which deflation is hurtful and variance increases.	85
5.2	Variance reduction ratios for matrices with spectra with linear, $\sigma_{N-i+1} = i$, quadratic, $\sigma_{N-i+1} = i^2$, and cubic, $\sigma_{N-i+1} = i^3$, growth rates.	85
5.3	On the left we have spectrum $\sigma_i = 1/\sqrt{i}$. The right plot shows the deflation of the inverse of a 2D discrete Laplacian on a grid $\sqrt{N} \times \sqrt{N}$ with Dirichlet boundary conditions.	86

5.4	Matrix BWM2000 has a size of $N = 2000$ and condition number of $2.37869e+5$. Matrix MHD1280B has $N = 1280$ and a condition number of $4.74959e+12$. Both matrices are real, non-symmetric.	88
5.5	Matrix NOS6 is symmetric, with $N = 675$ and a condition number of $7.65049e+06$. Matrix OLM1000 is non-symmetric, with $N = 1000$ and a condition number of $1.48722e+06$	89
5.6	Here the variance reduction of matrix BWM2000 and matrix MHD1280B is shown with Gaussian noise. We remind the reader that the sizes of these matrices are $N = 2000$ and $N = 1280$ respectively, with condition numbers of $2.37869e+5$ and $4.74959e+12$ correspondingly. Both matrices are real, non-symmetric.	90
5.7	NOS6 is $N = 675$ symmetric matrix, with and a condition number of $7.65049e+06$. OLM1000 is a $N = 1000$ non-symmetric matrix, with and a condition number of $1.48722e+06$. Both are deflated with Gaussian noise.	91
5.8	The variance reduction of matrices BWM2000 and MHD1280B is shown with Gaussian and \mathbb{Z}_4 noise. Both matrices are real and non-symmetric.	92
5.9	A comparison of variance reduction for matrices NOS6 and OLM1000 is shown. Since NOS6 is real and symmetric, its comparison is done with \mathbb{Z}_2 noise.	92
5.10	Above is the variance of the hierarchical probing trace estimator with and without deflation. The full 1000 vector subspace is used as the deflated operator in red. Complete color closings are marked with green circles. For the ensemble A matrix, a factor of 15 is achieved in variance reduction between deflated and undeflated probing. Deflation yields over a factor of 20 reduction of variance for the ensemble B matrix.	98
5.11	Speedup of the combined deflated HP estimator compared to pure \mathbb{Z}_4 noise is shown. The speedup to basic MC is estimated for both HP alone and HP with deflation. The errors are computed with Jackknife resampling.	99
5.12	The sum of squared absolute values of matrix elements at specified Manhattan distances from the corresponding diagonal elements for 10 randomly sampled rows. Base case is the original Monte Carlo method. Deflation refers to the Monte Carlo with deflation. HP and deflated HP refer to a space spanned by the 32 hierarchical probing vectors. A combination of HP and deflation suppresses the sum of matrix elements by orders of magnitude more than probing or deflation alone.	100
5.13	Variance for the ensemble A matrix as a function of the deflated SVD subspace dimension at two color closing points of HP. The left plot is with 32 probing vectors, the right is with the full 512.	101
5.14	Variance for the matrix from ensemble B, as a function of the deflated SVD subspace dimension at two color closing points of HP. The left plot is with 32 probing vectors, the right is with the full 512.	102

5.15	Eigenvectors computed by PRIMME from 100 to 1000 for the matrix from ensemble A. A log plot of variance and cost. Each case displays 5 points, which represent the variance and wallclock at probing vectors 32, 64, 128, 256, and 512.	103
6.1	A side by side comparison of the isoscalar disconnected axial charge at 32 and 64 HP vectors. There is a noticeable improvement scaling up to 64 HP vectors.	110
6.2	A side by side comparison of the isoscalar disconnected axial charge at 128 and 256 HP vectors. Error in the matrix element continues to decrease with up to 128 HP vectors. No significant improvement is observed with 256 HP vectors.	111
6.3	Above, the axial charge with deflated HP is compared to a pure noise trace estimator in the isoscalar sector. The points shown are with $t_{\text{sep}} = 10$, $\tau = 5$. Both methods are performed on the same 100 configurations.	112
6.4	A side by side comparison of the strange axial charge at 32 and 64 HP vectors. There is a noticeable reduction of error through 64 HP vectors. . .	112
6.5	A side by side comparison of the strange axial charge at 128, 256, and 512 HP vectors is shown. Error in the data appear lower with 128 HP vectors compared to 64. No significant improvement is observed with 256 or 512 HP vectors however.	113
6.6	A side by side comparison of the isoscalar disconnected scalar charge at 32 and 64 HP vectors. No difference is observed.	114
6.7	A side by side comparison of the isoscalar disconnected scalar charge at 128 and 256 HP vectors. No difference is observed.	115
6.8	The scalar charge with deflated HP is compared to a pure noise trace estimator in the isoscalar sector. The points shown are with $t_{\text{sep}} = 10$, $\tau = 5$. Both methods are performed on the same 100 configurations.	115
6.9	A side by side comparison of the strange scalar charge at 32 and 64 HP vectors. No difference is observed.	116
6.10	A side by side comparison of the strange scalar charge at 128, 256, and 512 HP vectors is plotted. Statistical errors are equivalent to even the most inaccurate computation of the trace in Figure 6.9a.	116
6.11	A side by side comparison of the isoscalar disconnected tensor charge at 32 and 64 HP vectors. Error reduces from 32 to 64 HP vectors.	117
6.12	A side by side comparison of the isoscalar disconnected tensor charge at 128 and 256 HP vectors. There is significant error reduction from 64 to 128 HP vectors. Additionally, the errors in the endpoints continue to decrease through 256 HP vectors.	118

6.13	The tensor charge with deflated HP is compared to a pure noise trace estimator in the isoscalar sector. The points shown are with $t_{\text{sep}} = 10$, $\tau = 5$. Both methods are performed on the same 100 configurations.	118
6.14	A side by side comparison of the strange tensor charge at 32 and 64 HP vectors. Errors, especially of the endpoints, reduce at 64 HP vectors compared to 32.	119
6.15	A side by side comparison of the strange tensor charge at 128, 256, and 512 HP vectors is shown. We observe only a little improvement from 64 to 128 HP vectors for the strange contribution. No clear gain is achieved beyond 128 HP vectors.	119
6.16	A side by side comparison of the isoscalar disconnected vector charge at 32 and 64 HP vectors. The statistical error decreases significantly from 32 to 64 HP vectors.	120
6.17	A side by side comparison of the isoscalar disconnected vector charge at 128 and 256 HP vectors. Errors continue to be reduced and the expected answer is approached.	121
6.18	The vector charge with deflated HP is compared to a pure noise trace estimator in the isoscalar sector. The points shown are with $t_{\text{sep}} = 10$, $\tau = 5$. Both methods are performed on the same 100 configurations.	122
6.19	A side by side comparison of the strange vector charge at 32 and 64 HP vectors. Errors reduce from 32 to 64 HP vectors.	123
6.20	A side by side comparison of the strange vector charge at 128, 256, and 512 HP vectors is given. We remark that the statistical error continues to go down, all the way through 512 HP vectors.	123
6.21	A correlated two-state model fit of the disconnected light axial a094m280 data is taken. The fit is with a t_{sep} range of 3-11. The $\chi^2/\text{d.o.f.}$ is 1.34. . .	126
6.22	The disconnected light axial data from the a127m285 ensemble are fitted with two different methods. The left plot is a fit to an exponential and the right is using the summation method. The exponential method has a source-sink separation fitting range of 6-12, while the summation fit employs a range of 5-12. The $\chi^2/\text{d.o.f.}$ are 1.37, and 1.43 respectively.	127
6.23	The strange axial a094m280 data points are fitted with three different functional forms. The left plot is an exponential fit, the middle is with the summation method, and the right is to the two-state model. The exponential fit has a t_{sep} range of 6-10, the summation method and two-state model have ranges of 3-10. The $\chi^2/\text{d.o.f.}$ are 1.42, 1.25, and 1.05 respectively. . .	128
6.24	The a127m285 strange axial data are fit to the two-state model from a t_{sep} range of 4-10. The fit is fully covariant. The $\chi^2/\text{d.o.f.}$ is 1.40.	129
6.25	We plot the a094m280 disconnected light data with two fits, one to an exponential and one employing the summation method. The left plot has a t_{sep} range of 12-14, right one a range of 9-14. The $\chi^2/\text{d.o.f.}$ are 1.33 and 1.47 for the respective left and right figures.	130

6.26	The a127m285 light scalar points are fitted with the two-state model. The fit is fully covariant. The source-sink separations included are 4-12. The $\chi^2/\text{d. o. f.} = 1.38$	131
6.27	The a094m280 strange scalar ratio is fitted to an exponential. The source-sink separations included are 4-7. The fit is uncorrelated and $\chi^2/\text{d. o. f.} = 0.94$	132
6.28	The a127m285 light scalar points are fitted with the two-state model. The fit is fully covariant. The source-sink separations included are 4-12. The $\chi^2/\text{d. o. f.} = 1.38$	132
6.29	The ratio of g_S for the proton-proton divided by the proton is plotted. The proton t_{sep} is fixed at 5-8 for the exponential data and 6-8 for the summation method and two-state model results. The t_{sep} ranges of pp follow from Tables 6.4, 6.6, and 6.8.	139
6.30	The ratio of g_S for the deuteron to proton ratio is given. The proton t_{sep} is fixed at 5-8 for the exponential data and 6-8 for the summation method and two-state model results. The t_{sep} ranges of the deuteron are the columns of Tables 6.10, 6.12, and 6.14.	142
6.31	The scalar charge of ${}^3\text{He}$ divided by the scalar charge of the proton is plotted. The proton t_{sep} is 4-8 for the exponential data and 6-8 for the summation method and two-state model. The t_{sep} ranges of the ${}^3\text{He}$ charge are the columns of Tables 6.16, 6.18, and 6.20.	145

DISCONNECTED DIAGRAMS IN LATTICE QCD

CHAPTER 1

Motivation: Hadron Structure

For nearly half a century, quantum chromodynamics (QCD) has been accepted as the theory for describing strong interactions, yet many difficulties remain in fully understanding hadron structure. At high energies, perturbative QCD has enjoyed tremendous success in agreeing with experimental results from PETRA in DESY, LEP in CERN, and many others. At low energies however, an analytical solution of QCD is not possible, resulting in an incomplete understanding of the precise mechanism in which quarks and gluons combine together to form the nucleon and other hadrons. A clear picture of the exact dynamics which gives rise to the mass, spin content, and charge distribution of the nucleon has yet to be achieved.

Numerous experiments have been planned or are already ongoing with the purpose of enhancing our view of hadron structure. To name a few, COMPASS at CERN is investigating the origin of proton spin and exotics. The nucleon form factors at high momentum transfer will be probed at CEBAF in Jefferson Lab under the 12 GeV upgrade. Additionally, the GlueX project in Jefferson Lab is mapping the spectrum of potential exotic mesons. MAMI in MAINZ is studying nucleon form factors and polarizabilities.

To complement these experimental efforts, a rigorous treatment of QCD at low energies is required to make predictions or confirm experimental results. The only model independent and systematically improvable way to theoretically study QCD at low energies is Lattice QCD. This entails numerically computing the Feynman path integral on a lattice stochastically with Monte Carlo methods. In order to do this, the time dimension is analytically continued to the imaginary axis of the complex plane, resulting in a theory with Euclidean space-time. Euclidean correlators on the lattice may then be computed as

$$\langle \mathcal{O}_2(t)\mathcal{O}_1(0) \rangle = \frac{1}{Z} \int \mathcal{D}[\psi, \bar{\psi}] \mathcal{D}[U] e^{-S_F[\psi, \bar{\psi}, U] - S_G[U]} \mathcal{O}_2[\psi, \bar{\psi}, U](t) \mathcal{O}_1[\psi, \bar{\psi}, U](0), \quad (1.1)$$

where Z is the partition function, described by

$$Z = \int \mathcal{D}[\psi, \bar{\psi}] \mathcal{D}[U] e^{-S_F[\psi, \bar{\psi}, U] - S_G[U]}. \quad (1.2)$$

Details of the definitions of ψ , U , $S_G[U]$, and $S_F[\psi, \bar{\psi}, U]$ are given in Chapter 2. Lattice QCD simulations have limitations such as finite lattice spacing, finite volume, and unphysical light quark masses, typically labeled by an unphysical pion mass. The lattice spacing is used as a non-perturbative regulator, whereas an unphysical pion mass is because of computational restrictions. Due to improved algorithms and hardware however, computations at the physical point have started to appear in recent years [1, 2, 3]. A full Lattice calculation involves computing observables on different ensembles and extrapolating to zero lattice spacing, infinite volume, and if needed, physical pion mass.

The computational expense of Lattice calculations grows as the box volume increases and as the quark mass decreases, this will be discussed further in Chapters 4 and 5. As computers become more powerful, Lattice QCD calculations are able to access finer

lattice spacings, larger volumes, and physical quark masses. Unfortunately, the error from averaging over an ensemble of gauge field configurations scales as a typical Monte Carlo process: $\frac{1}{\sqrt{N}}$, where N is the number of measurements. Therefore, a factor of 100 increase in computing capability is necessary to reduce the error by an order of magnitude. In order for the field to truly prosper, more efficient computers must be combined with smarter algorithms that reduce the variance/error of the computed observables.

Correlation functions in Lattice QCD are broadly defined into two categories: connected and disconnected pieces. These classifications arise from the fermion Wick contractions of hadrons and will be rigorously defined in Section 3.2. We define disconnected contributions here loosely, as an operator insertion in a quark loop coupled to a hadron via only gluons. Over the past decade, most studies of nucleon observables either neglect disconnected diagrams or limit their study to cases in which these diagrams cancel, such as the isovector channel. This is due to the notoriously high computational cost of computing a quark loop on the lattice. In this thesis we study state-of-the-art methods to compute disconnected diagrams and their applications.

The remainder of this chapter covers three interesting areas of physics in which evaluation of disconnected diagrams can give appreciable insight. Chapter 2 briefly reviews continuum QCD and introduces the reader to Lattice QCD and the general lattice action that is used in all computations in this work. Chapter 3 describes how nucleon matrix elements are computed on the lattice. Chapter 4 is an introduction to Numerical Linear Algebra and the basic methods that many collaborations employ to calculate disconnected diagrams. Additionally, a succinct overview of hierarchical probing (HP) and its benefits is given at the end of Chapter 4. The first half of Chapter 5 describes the details of SVD deflation and provides a model that qualitatively predicts the effects of deflation on various matrices from a plethora of fields, including chemical transport modeling, magnetohydrodynamics, and electrical engineering. The second segment of Chapter 5 focuses on the

direct application of SVD deflation to disconnected diagrams in Lattice QCD and notes the synergy between HP and deflation. Chapter 6 summarizes results and applications of the previous chapters to nucleon and light nuclei matrix elements. The disconnected axial, scalar, tensor, and vector charges are shown for the nucleon as well as sigma terms for the proton, proton-proton system, deuteron, and ${}^3\text{He}$.

1.1 Proton spin crisis

Spin is one of the fundamental properties of a particle, along with its mass and charge. It governs a particle's statistical properties and equations of motion. Spin is relevant to almost all areas of physics, from quantum computing to condensed matter theory. The drastic temperature difference in achieving ${}^3\text{He}$ superfluidity versus ${}^4\text{He}$ superfluidity is just one of many phenomena deeply connected with the physics of spin. Therefore, understanding the spin of the nucleon is of central importance.

The proton spin crisis is one of the modern-day challenges of nuclear physics. Although the proton is a spin-1/2 particle, akin to all the quarks, the precise contributions to its spin are still unknown. The story begins in 1987 due to an experiment by the European Muon-Collaboration (EMC). Before this experiment it was widely believed that the entirety of a proton's spin came from its constituent quarks. This simple picture constrained two of its quarks to carry spin in opposite directions, with the third quark's spin aligned in the same direction as the proton's spin.

In the EMC experiment, a polarized muon beam was incident on a polarized proton target, measuring the instantaneous spin of a quark. The experimental data indicated that the sum of quark spins inside the proton was only a fraction of its total spin. These results were confirmed in numerous subsequent experiments. The issue of where the missing spin comes from became known as the proton spin crisis.

It's now believed that the missing spin is carried by a combination of gluon spin, quark orbital angular momentum, and gluon orbital angular momentum.

$$\frac{1}{2} = \frac{1}{2}\Sigma_q + \Sigma_g + L_q + L_g \quad (1.3)$$

All contributions to the nucleon's spin are written above, here \hbar has been set to 1. There are subtleties in uniquely defining angular momentum for the quarks and gluons which we do not cover here. We may define the proton's spin through the zero-momentum transfer matrix element of the axial-vector current (the axial charge).

$$S_\mu = \frac{1}{2m_P} \langle p, s | \bar{u}_P(p, s) \gamma_\mu \gamma_5 u_P(p, s) | p, s \rangle \quad (1.4)$$

Above, u_P , \bar{u}_P are the proton spinors and m_P is the proton mass. Individual axial charges for different quark flavors can then be interpreted as the quark contribution to the total proton spin.

$$\Delta q S_\mu = \frac{1}{2m_P} \langle p, s | \bar{q} \gamma_\mu \gamma_5 q | p, s \rangle \quad (1.5)$$

These up, down, and strange axial charges can be written in combinations of the isovector, octet, and flavor-singlet axial charges.

$$g_A^{(3)} = \Delta u - \Delta d \quad (1.6)$$

$$g_A^{(8)} = \Delta u + \Delta d - 2\Delta s \quad (1.7)$$

$$g_A^{(0)} = \Delta u + \Delta d + \Delta s \quad (1.8)$$

Equation 1.8 gives the full quark contribution to the proton spin. One may estimate

the size of these charges with relativistic quark confined models, one such example is the MIT bag model.

In the MIT bag model, three non-interacting quarks are confined to a spherical cavity, or elastic “bag”. If one attempts to pull one of the quarks out, this bag stretches, giving rise to a linear confining potential energy. The amount of energy required to overcome this force and free a quark is far above the quark-antiquark pair production threshold, resulting in a jet of mesons. Computing the flavor-singlet axial charge with this model results in the quarks carrying roughly 60% of the spin of the proton.

From polarized deep inelastic scattering (pDIS) experiments however, the flavor-singlet charge has been determined to only contribute approximately 30% to the proton spin. Furthermore, the SU(3) octet axial charge may be extracted from hyperon beta-decay [4], allowing the combination of both charges to isolate the strange spin content of the proton. From fitting with modern models to experimental results [5], the strange axial charge is currently quoted to be $\Delta s = -.03 \pm 03$ in the MS-bar scheme at a scale of 2 GeV.

An ab initio calculation of the nucleon axial charge is crucial in comparing with experimental results and furthering our understanding of how QCD dynamics plays a role in forming hadrons. In Section 3.2 the different Wick contractions of nucleon matrix elements will be discussed, resulting in two types of diagrams: connected and disconnected. Both categories of diagrams are important in understanding the role quarks play in contributing to the proton spin. As will be shown in Chapter 3, the strange spin content of the proton is an effect solely due to disconnected diagrams.

1.2 Dark matter cross sections

It is well known that the luminous mass density of spiral galaxies decreases as a function of distance from the center. Therefore, if luminous mass was the only type of matter present, rotation curves of spiral galaxies would be expected to decrease with distance. However, in 1939 Horace W. Babcock observed a relatively flat rotation curve for the Andromeda nebula [6]. Since then, the behavior of rotation curves from many other galaxies have matched his results [7]. The simplest way to account for this is to assume the existence of non-luminous (“dark”) matter, which forms halos around the galactic disks.

Other indications of dark matter include velocity dispersions of elliptical galaxies, gravitational lensing, anisotropies in cosmic microwave background radiation, and many more observations [8, 9]. In all cases, introducing dark matter can be used to explain the difference between observed and expected phenomena. Utilizing these constraints allows for a prediction of the precise mass-energy budget for the known universe. Ordinary and visible baryonic matter only makes up about 5% of the total, with dark matter accounting for $\approx 25\%$, and the remaining 70% attributed to dark energy, a form of energy postulated to explain the expansion of the universe.

The most popular dark matter candidates are weakly interacting massive particles (WIMPs). Based on the amount of dark matter required to explain the observations listed above, the mass scale of WIMPs is around 100 GeV. This mass scale is within range of LHC experiments that are currently underway. In order to write down the general spin-independent interaction of a nucleus with a WIMP, the scalar matrix elements (sigma terms) are needed.

$$\sigma_l = \frac{m_u + m_d}{2} \langle N | \bar{u}u + \bar{d}d | N \rangle \quad (1.9)$$

$$\sigma_s = m_s \langle N | \bar{s}s | N \rangle \quad (1.10)$$

Above, σ_l is the light sigma term and σ_s is the strange sigma term; the N state can refer to either a single nucleon or a nucleus. These sigma terms are the main source of uncertainty in calculating dark matter cross sections [10, 11, 12].

Sigma terms cannot be accessed directly, however there is a procedure to relate them to pion-nucleon scattering. The light sigma term can be obtained through measuring the isoscalar πN scattering amplitude with the pseudovector Born terms subtracted out ($\Sigma_{\pi N}$). In order for σ_l and $\Sigma_{\pi N}$ to match up to order $\mathcal{O}(m_l^2)$, the πN amplitude is analytically continued to an unphysical kinematic region, the Cheng-Dashen point (defined by the Mandelstam variables $u = s = m_N^2, t = m_\pi^2$).

The difference between the pion-nucleon scattering amplitude and the scalar form factor at the Cheng-Dashen point, Δ_R , was computed using heavy baryon chiral perturbation theory in [13]. Δ_R was found to have an upper limit of about 2 MeV. The zero-momentum transfer scalar matrix element could then be extracted and was found to be roughly 79(7) MeV in [14]. A recent alternative analysis using a combination of Lorentz covariant baryon chiral perturbation theory and Hellmann-Feynman theorem gives $\sigma_l = 59(7)$ MeV [15].

The strange sigma term is more difficult to estimate due to the strange quark's much heavier mass, which changes the leading order matching of the scattering amplitude and scalar form factor. Instead, the octet baryon mass splittings are used to estimate the flavor-singlet sigma term.

$$\sigma_0 = \frac{m_u + m_d}{2} \langle N | \bar{u}u + \bar{d}d - 2\bar{s}s | N \rangle \quad (1.11)$$

An analysis utilizing this method in [16] yielded $\sigma_0 = 36(7)$ MeV. The strange sigma term is then related with

$$\sigma_s = \frac{m_s}{m_u + m_d}(\sigma_l - \sigma_0). \quad (1.12)$$

Unfortunately, due to the large numerical factor in the front of (1.12), this estimation leads to large uncertainties. Due to these difficulties, there is tremendous opportunity for Lattice QCD to precisely determine this matrix element. Identical to the strange proton spin content, the strange scalar matrix element arises solely due to the presence of disconnected diagrams in the Wick contractions of the scalar three-point correlation function. Recent Lattice QCD studies have directly computed pion and nucleon sigma terms [17, 18], however before this work we are not aware of any results on multi-hadronic states. This is of utmost importance, as dark matter detection experiments use Germanium, liquid argon, and liquid xenon for the detector media [19]. Furthermore, from the EMC effect [20], it is well known that properties of nuclei differ from the sum of free nucleons. Therefore, a calculation of sigma terms of light nuclei to an accuracy which resolves the difference from naive linear scaling serves as invaluable input for dark matter detection experiments.

1.3 Proton radius puzzle

The charge radius of the proton, or any atomic nucleus, is a quantification of its size. Of course neither the proton nor other nuclei have concrete boundaries, resulting in an electron experiencing a range of “radii” in scattering experiments. The charge radius is defined as the root mean square of these radii. To formally define it, we introduce the nucleon vector current matrix element.

$$\langle p' | \bar{q} \gamma_\mu q | p \rangle = \bar{u}(p') \left[\gamma_\mu F_1^q(Q^2) + \frac{i}{2M_P} \sigma_{\mu\nu} q^\nu F_2^q(Q^2) \right] u(p) \quad (1.13)$$

Above, $\sigma_{\mu\nu} = \frac{i}{2} [\gamma_\mu, \gamma_\nu]$, $q = p' - p$ is the momentum-transfer, $Q^2 = -(p' - p)^2$, and $F_1^q(Q^2)$, $F_2^q(Q^2)$ are the flavor-dependent Pauli and Dirac form factors respectively. This matrix element characterizes the electromagnetic interaction vertex of a photon being absorbed by a proton. A linear combination of $F_1(Q^2)$ and $F_2(Q^2)$ can be taken to form the electric and magnetic form factors.

$$G_E(Q^2) = F_1(Q^2) - \frac{Q^2}{4M_P^2} F_2(Q^2) \quad (1.14)$$

$$G_M(Q^2) = F_1(Q^2) + F_2(Q^2) \quad (1.15)$$

In the nonrelativistic limit, $G_E(Q^2)$ and $G_M(Q^2)$ are simply Fourier transforms of the nucleon charge and current densities. Moreover, the electric and magnetic form factors carry information about the extended nature of a proton and describe its deviation from a point-like object.

Now a formal definition of the electric and magnetic radii can be given.

$$r_E^2 = -6 \left. \frac{dG_E(Q^2)}{dQ^2} \right|_{Q^2=0} \quad (1.16)$$

$$r_M^2 = -6 \left. \frac{dG_M(Q^2)}{dQ^2} \right|_{Q^2=0} \quad (1.17)$$

There are two methods for obtaining the proton charge radius via electrons. The first is through fitting the electric and magnetic form factors with elastic electron-proton scattering cross sections [21]. The second is to measure the Lamb shift in ordinary electronic

hydrogen [22]. There are subtleties in the second technique involving uncertainty with the Rydberg constant that we do not cover here. Both experiments are compatible with a value of $r_E = .88(01)$ fm.

Another method for computing the proton charge radius from experiment is to study the Lamb shift of muonic Hydrogen, instead of ordinary Hydrogen. The muon is roughly 200 times heavier than the electron, so it “orbits” 200 times closer to the proton in the bound system. This allows the radius to be determined with extreme accuracy. Recent results from muonic hydrogen give $r_E = .84087(39)$ fm [23]. This is 4% smaller than the electronic measurements, with an extremely small uncertainty. There is a 7σ difference between the muonic hydrogen measurement and previous results. Even more recently, the charge radius of the deuteron was measured with muonic deuterium [24]. The uncertainty in the deuteron radius is appreciably smaller than any previous measurements, however this measured radius gives a 7σ discrepancy from the expected value.

1.3.1 Strange radius

As an important aside, through parity-violating electron scattering the neutral weak current form factors of the proton can be measured. There are two vector current form factors F_1^Z , F_2^Z , and an axial-vector current form factor G_A^Z that describe this interaction. Combining the two vector form factors into electric/magnetic pieces (similar to the previous discussion) and assuming isospin symmetry gives a clean relationship between the photon and weak vector form factors.

$$G_E^{Z,p} = (1 - 4 \sin^2 \theta_W) G_E^p - G_E^n - G_E^s \quad (1.18)$$

$$G_M^{Z,p} = (1 - 4 \sin^2 \theta_W) G_M^p - G_M^n - G_M^s \quad (1.19)$$

Above, the form factors on the right hand side are the vector-current ones defined in (1.14) and (1.15). This relation allows the strange vector current form factor to be isolated and conversely the strange radius to be computed with (1.16). Experiments at Jefferson Lab have found the strange vector current form factor of the proton to be consistent with zero [25].

There is great scope for Lattice QCD calculations to aid in the effort of understanding the charge radius puzzle. A high precision computation of the strange form factors and radii from Lattice QCD was conducted in [26]. The study yielded an extremely precise non-zero value for the strange radius, which agreed with results from Jefferson Lab due to a larger uncertainty on the experimental side. In order to have systematics fully under control however, this calculation must be repeated at multiple lattice spacings and an extrapolation to the continuum must be taken. Additionally, the light quark disconnected diagrams are also necessary [27].

We have only given a basic survey of the very extensive topics discussed here. For further details, we refer the reader to [28] and [29] for information on the charge radius puzzle and parity-violating electron scattering respectively.

CHAPTER 2

Introduction to Lattice QCD

In this chapter an overview of quantum chromodynamics (QCD) on the lattice is presented. Basic knowledge of quantum field theory is assumed. Only the basic methodology is described in this segment and we refer the interested reader to [30] and [31] for further details.

QCD is the theory of quarks and gluons, the strongly interacting particles. We begin by reviewing the continuum QCD action functional and its symmetries in the first section. This section is ended with a brief discussion of asymptotic freedom and the need for non-perturbative methods such as Lattice QCD at low energies. We follow with a discretization of the gluonic part of the QCD action and describe the Wilson pure gauge lattice action. In the subsequent section, a discretization of the fermionic part in the most straight forward and naive way is discussed. This naturally leads to the problem of Fermion doubling, which we briefly summarize, and we propose the Wilson Fermion action as one possible solution to this doubling. This chapter is concluded with an examination of discretization errors and the introduction of the Clover term.

2.1 Continuum QCD

This review is started with a description of the quark and gluon fields for the continuum Euclidean action. Subsequently, we examine the fermionic and gluonic pieces of the action independently. By studying the gauge invariance of the fermionic part, the covariant gauge derivative is introduced in a natural context. We may then verify the invariance of the gluonic piece under the same transformations.

2.1.1 Fields in the QCD action

There are two types of fundamental fields in the theory of QCD, whose excitations produce quarks and gluons. Quarks are spin 1/2 fermions with 4-spinor Dirac structure, similar to fermions in quantum electrodynamics (QED). Dirac spinors are irreducible representations of the Poincaré group. A major difference between QCD and QED, however, is that quarks also carry a color index.

$$\psi^{(f)}(x)_{\alpha,c} \tag{2.1}$$

Here, x is the space-time coordinate, α is the Dirac index, c is the color index, and f is the flavor index. Since $\alpha = 1, 2, 3, 4$ and $c = 1, 2, 3$, for each flavor, $\psi^{(f)}$ has 12 independent components. The color index arises from SU(3) gauge symmetry, in contrast to the U(1) symmetry from QED. There are six different flavors of quarks: up, down, strange, charm, bottom, and top. We will often drop indices in lieu of matrix/vector notation. Apart from being notationally convenient, this also makes direct application of large matrix techniques from numerical linear algebra obvious. This will be discussed further in Chapters 4 and 5.

With a Minkowskian action, one may define $\bar{\psi} = \psi^\dagger \gamma_0$, where γ_0 is the time γ matrix. For the Euclidean action, no such relation exists; ψ and $\bar{\psi}$ are treated as separate integra-

tion variables. The other type of field in QCD is the gluon field, corresponding to a spin 1 vector gauge boson.

$$A_\mu(x)_{cd} \tag{2.2}$$

Once again, x is the space-time coordinate. Additionally, these gauge fields carry a Lorentz index: μ , which corresponds to different space-time components. In the Euclidean case, there is no need to discriminate between contravariant and covariant indices since the metric tensor is the identity. The gluon fields have two color indices: c, d . Specifying a space-time position (x), and direction (μ), yields a 3×3 traceless, hermitian matrix for $A_\mu(x)$. This comes from the fact that $A_\mu(x)$ may be written as a linear combination of the generators of SU(3). In the following subsections we inspect the fermionic and gluonic parts of the QCD action individually.

2.1.2 The fermionic QCD action

It should be noted that even the fermionic piece of the QCD action depends on the gauge field. This is due to the coupling between quarks and gluons, which arises to enforce gauge invariance (see Section 2.1.3). The fermionic part with a Euclidean metric, $S_F[\psi, \bar{\psi}, A]$ is given as

$$S_F[\psi, \bar{\psi}, A] = \sum_{f=1}^{N_f} \int d^4x \bar{\psi}^{(f)}(x)_{\alpha,c} [(\gamma_\mu)_{\alpha\beta} (\delta_{cd}\partial_\mu + iA_\mu(x)_{cd}) + m^{(f)}\delta_{\alpha\beta}\delta_{cd}] \psi^{(f)}(x)_{\beta,d}. \tag{2.3}$$

Above, we use Greek letters and ordinary lower case letters to differentiate between Dirac and color indices, respectively. Additionally, we leave the sum over flavors as running from 1 to N_f . Although N_f is 6 in the full theory, for calculations at low energies,

only including the lightest flavors is adequate. At energy ranges of the proton mass, the contribution from the heavier 3 flavors is expected to be small and their effect can be tucked into the effective parameters of the theory. From (2.3), it can be inferred that different quark flavors couple to the color gauge field identically, and the only dissimilarity arises from their masses: $m^{(f)}$.

The quark fields, ψ and $\bar{\psi}$, couple to the gauge field through the contraction of the color indices c and d . Additionally, the gauge field contracts with various γ matrices. The γ matrices in (2.3) are the Euclidean analogs of the Minkowski γ matrices that appear in the Dirac equation. In Appendix A, an explicit description of these γ matrices is given. Similar to the gauge field, the kinetic term of (2.3) also contracts with γ_μ , and also with the quarks through the Dirac indices α and β . The kinetic term is trivial in the color space and the mass term is trivial in both Dirac and color spaces.

Using the Euler-Lagrange equations, we can derive the classical equation of motion for this action.

$$\partial_\mu \left(\frac{\partial L}{\partial(\partial_\mu \bar{\psi})} \right) - \frac{\partial L}{\partial \bar{\psi}} = 0 \quad (2.4)$$

Applying this to (2.3) produces

$$\sum_{f=1}^{N_f} [\gamma_\mu (\partial_\mu + iA_\mu(x)) + m^{(f)}] \psi^{(f)}(x) = 0. \quad (2.5)$$

This is the Euclidean Dirac equation in an external gauge field, A_μ .

2.1.3 Verifying Gauge invariance

Now that the basic form of the fermionic QCD action has been presented we may study its symmetries. It should be noted that apart from the color structure, (2.3) is

identical to the fermionic QED action. Furthermore, generalizing the gauge invariance of QED inherently leads to the QCD action.

The QED action is invariant under the multiplication of an arbitrary phase at each space-time point, conjoined with a gauge field transformation of the photon field. In QCD, this invariance must be under local rotations in the color space of the quark fields. At each space-time point we introduce 3×3 unitary matrices $\Omega(x)$, which are group elements of $SU(3)$. Note that these elements do not commute in general, therefore the group is non-abelian.

We may impose gauge invariance of QCD with these $SU(3)$ matrices. For the quark fields, this takes the form

$$\begin{aligned}\psi'(x) &= \Omega(x)\psi(x) \\ \bar{\psi}'(x) &= \bar{\psi}(x)\Omega(x)^\dagger,\end{aligned}\tag{2.6}$$

where $\psi'(x)$ and $\bar{\psi}'(x)$ are the fermion fields after this gauge transformation. The gauge field $A_\mu(x)$ must also transform in some way to uphold this invariance. Applying this transformation with (2.3), we arrive at

$$S_F[\psi', \bar{\psi}', A'] = \int d^4x \bar{\psi}(x)\Omega(x)^\dagger \left[\gamma_\mu \left(\partial_\mu + iA'_\mu(x) \right) + m \right] \Omega(x)\psi(x),\tag{2.7}$$

where the flavor sum, Dirac, and color indices have been suppressed. Due to $\Omega(x)^\dagger = \Omega(x)^{-1}$, the mass term is trivially invariant under this transformation. The other terms are not so obvious, comparing (2.3) and (2.7) yields

$$\partial_\mu + iA_\mu(x) = \Omega(x)^\dagger \left[\partial_\mu + iA'_\mu(x) \right] \Omega(x) \quad (2.8)$$

$$= \partial_\mu + \Omega(x)^\dagger (\partial_\mu \Omega(x)) + i\Omega(x)^\dagger A'_\mu(x) \Omega(x). \quad (2.9)$$

Now we are able to write a transformation rule for the gauge field.

$$A'_\mu(x) = \Omega(x) A_\mu(x) \Omega(x)^\dagger + i (\partial_\mu \Omega(x)) \Omega(x)^\dagger \quad (2.10)$$

With the transformation from (2.10), we have all the pieces in place to investigate the gluonic part of the action.

2.1.4 The gluonic QCD action

We begin this discussion by introducing the covariant gauge derivative.

$$D_\mu(x) = \partial_\mu + iA_\mu(x) \quad (2.11)$$

We define this new derivative by inspecting (2.8) and noting that

$$D'_\mu(x) = \partial_\mu + iA'_\mu(x) = \Omega(x) D_\mu(x) \Omega(x)^\dagger, \quad (2.12)$$

which guarantees that $D_\mu(x)\psi(x)$ and $\psi(x)$ transform identically under a gauge rotation.

We may construct a gluonic action using these covariant gauge derivatives, which is an extension of the expression from QED that depends on the field strength tensor, $F_{\mu\nu}(x)$.

We define this tensor as the commutator of the covariant derivative, akin to QED.

$$F_{\mu\nu}(x) = -i [D_\mu(x), D_\nu(x)] = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) + i [A_\mu(x), A_\nu(x)] \quad (2.13)$$

In QED the commutator over the gauge fields vanishes, however in QCD $A_\mu(x)$ and $A_\nu(x)$ are matrices, and their commutator is in general non-zero. This leads to cubic and quartic gluon interactions seen in Figure 2.1.

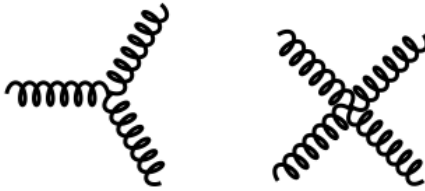


FIG. 2.1: Feynman diagrams representing the gluon self interaction terms found in the QCD action. Unlike photons in QED, gluons carry color charge, which engenders the non-linear nature of the strong force.

It may also easily be checked that the field strength tensor transforms as in (2.12).

$$F'_{\mu\nu}(x) = \Omega(x)F_{\mu\nu}(x)\Omega(x)^\dagger \quad (2.14)$$

We may extend the consonance with QED further, and write down the gluonic QCD action as

$$S_G[A] = \frac{1}{2g^2} \int d^4x \text{Tr} [F_{\mu\nu}(x)F_{\mu\nu}(x)]. \quad (2.15)$$

As a reminder, due to our Euclidean prescription, covariant and contravariant forms do not need to be distinguished. The contraction in Lorentz indices, μ and ν , ensures the action is a Lorentz scalar. The astute reader will note that unlike in QED, the QCD action contains a trace. This trace acts as a reduction operator since the gluon fields are

matrices. Additionally, from (2.14) it may be verified that this action is gauge-invariant due to the cyclic property of the trace. Furthermore, this term is a dimension 4 operator. Therefore, this expression meets the requirements of Lorentz invariance, gauge invariance, and renormalizability.

2.1.5 Asymptotic freedom

Just as in QED, the coupling parameter of QCD “runs” with the renormalization scale. The bare coupling is of course independent of the renormalization scale. The running of this coupling parameter is described by the beta function $\beta(g)$.

$$\beta(g) = \frac{\partial g}{\partial \log(\mu)} \quad (2.16)$$

The beta function of QCD, computed from one-loop perturbation theory is

$$\beta(\alpha_s) = \frac{\alpha_s^2}{\pi} \left(\frac{N_f}{3} - \frac{33}{6} \right), \quad (2.17)$$

where $\alpha_s = \frac{g^2}{4\pi}$. This function is negative as long as $N_f < 17$, giving rise to the property of decreasing coupling parameter at increasing energy, or renormalization scale. This may be understood from the three and four-vertex gluon interactions causing an anti-screening effect. This feature is known as asymptotic freedom and it has allowed tremendous success of the parton model and collinear factorization to explore scattering processes [32]. Conversely, at low energies the coupling parameter becomes larger. At the mass scale of hadrons such as the pion or proton, this coupling is large enough to cause perturbation theory to break down. Therefore a non-perturbative treatment of QCD is required to study the properties of hadrons from the fundamental theory. We achieve this by discretizing QCD onto a lattice, giving the role of the lattice spacing as a non-perturbative regulator.

2.2 The Wilson pure gauge action

In this section we convert the gluon fields to link variables as the fundamental objects to use on the lattice. We perform a discretization and write down the Wilson gauge action. Finally, we demonstrate that in the limit of lattice spacing going to zero the continuum theory is recovered.

2.2.1 Gauge fields on the lattice as link variables

The first step in developing a discretization of a continuum theory is to introduce a lattice. In our case this is a 4D lattice.

$$\Lambda = \{(n_1, n_2, n_3, n_4)\} \tag{2.18}$$

One can interpret n as a vector on the lattice, $n \in \Lambda$, that can take on the value of space-time points separated by the lattice spacing a . In Lattice QCD, typically space and time directions have different total box lengths, so we differentiate between these two maximum values as N_L and N_T .

Let us consider an object that lives on this lattice: $U_\mu(n)$. It contains both a directional index μ and a lattice coordinate, as such $U_\mu(n)$ may be understood as a type of lattice vector field. Veritably, this object is matrix valued, causing its structure to be more complex. Let us postulate that the gauge transformation of this object yields

$$U'_\mu(n) = \Omega(n)U_\mu(n)\Omega(n + \hat{\mu})^\dagger. \tag{2.19}$$

Above, $\Omega(n)$ represents an element of $SU(3)$ for each lattice site, this is a discretized version of the color rotation from the continuum theory. The $\hat{\mu}$ is the unit vector in the μ direction; $n + \hat{\mu}$ is therefore one hop in the μ direction. We leave the verification of the

above gauge transformation to the next section.

Since there is a $U_\mu(n)$ corresponding to each direction μ , it is natural to view these objects as “link variables”. For example, $U_\mu(n)$ is the link that hops from sites n to $n + \hat{\mu}$. A link that goes in the opposite direction may also be defined: $U_{-\mu}(n)$, which hops from n back to $n - \hat{\mu}$. These two link variables are related by the property

$$U_{-\mu}(n) = U_\mu(n - \hat{\mu})^\dagger. \quad (2.20)$$

These links are elements of $SU(3)$ and are the gluon fields on the lattice. We will verify this, and give further interpretation of these links in the continuum in the next subsection.

2.2.2 Link variables in the continuum

Let us now relate these link variables to the gauge fields from continuum QCD. In the previous section we postulated $U_\mu(n)$ transforms as (2.19) under a gauge transformation. In the continuum, an object that transforms this way is known as a gauge transporter. Gauge transporters are path-ordered exponentials of the gauge field A_μ , integrated along some curve from points x to y .

$$G(x, y) = P e^{i \int_{curve} A \cdot ds} \quad (2.21)$$

Under a local symmetry rotation, these objects transform as

$$G'(x, y) = \Omega(x) G(x, y) \Omega(y)^\dagger. \quad (2.22)$$

From the matching transformation properties, we identify link variables as lattice

counterparts to the continuum gauge transporters. On the lattice, $U_\mu(n)$ connects points n and $n + \hat{\mu}$, therefore

$$U_\mu(n) = e^{iaA_\mu(n)}. \quad (2.23)$$

2.2.3 Gauge-invariant objects on the lattice

Now that gauge links have been distinguished as the fundamental gluonic degrees of freedom on the lattice, we may begin to formulate the gluon action. As a prerequisite, it is important to ascertain what gauge invariant objects can be constructed from link variables. The gluon action must be gauge-invariant, so the first step in writing down a sensible gluon action will be to discover functions of the lattice gauge transporters that are locally invariant under color rotations.

Let us begin by defining $P[U]$, an ordered product of N links from lattice sites n_1 to n_2 .

$$P[U] = U_{\mu_1}(n_1)U_{\mu_2}(n_1 + \hat{\mu}_1)\dots U_{\mu_N}(n_2 - \hat{\mu}_N) \quad (2.24)$$

Note that $P[U]$ may have forward or backward links, depending on the path. From (2.19), it is trivial to show that all intermediate rotations cancel and only the end point color rotation matrices remain. Therefore $P[U]$ transforms analogously to a single link variable.

$$P'[U] = \Omega(n_1)P[U]\Omega(n_2)^\dagger \quad (2.25)$$

Let us further define a specific class of paths $L[U]$ as closed paths $P[U]$ which begin and end at the same lattice site. If the trace of such a quantity is taken, it follows from

the cyclic property that the beginning and ending color rotation matrices will also cancel.

$$\text{Tr}(L'[U]) = \text{Tr}(\Omega(n_1)L[U]\Omega(n_1)^\dagger) = \text{Tr}(L[U]) \quad (2.26)$$

Now that we have established that the trace of closed loops of links are gauge-invariant quantities on the lattice, we ready to write down the gluon action.

2.2.4 The Wilson gauge action

The smallest non-trivial gauge-invariant object is called the plaquette. It is defined as a closed square with a length of one gauge link, or

$$U_{\mu\nu}(n) = U_\mu(n)U_\nu(n + \hat{\mu})U_{-\mu}(n + \hat{\mu} + \hat{\nu})U_{-\nu}(n + \hat{\nu}). \quad (2.27)$$

Using (2.20), this expression can be rewritten as

$$U_{\mu\nu}(n) = U_\mu(n)U_\nu(n + \hat{\mu})U_\mu(n + \hat{\mu})^\dagger U_\nu(n)^\dagger. \quad (2.28)$$

Applying the result of the previous subsection, we find that the trace of the plaquette is a gauge-invariant quantity. Therefore, the trace of the plaquette is the simplest candidate for a gluonic lattice action. Indeed, the Wilson gauge action is a sum over the trace of all plaquettes on the lattice, and we will show that in the limit of zero lattice spacing the continuum QCD gluon action is recovered. This sum is over all lattice points n , and all directions $1 \leq \mu < \nu \leq 4$, where these limits are placed to prevent double counting.

$$S_G[U] = \frac{2}{g^2} \sum_{n \in \Lambda} \sum_{\mu < \nu} \text{ReTr}[I - U_{\mu\nu}(n)] \quad (2.29)$$

I is the 3×3 identity matrix and the factor of $\frac{2}{g^2}$ ensures that the correct expression

is reached in the continuum limit. To demonstrate the equivalence between the Wilson gauge action in the limit of zero lattice spacing and the gluon QCD action, we begin by expanding the plaquette with (2.23).

$$\begin{aligned}
U_{\mu\nu}(n) \approx & \exp \left(iaA_\mu(n) + iaA_\nu(n + \hat{\mu}) - iaA_\mu(n + \hat{\nu}) - iaA_\nu(n) - \frac{a^2}{2} [A_\mu(n), A_\nu(n + \hat{\mu})] \right. \\
& - \frac{a^2}{2} [A_\mu(n + \hat{\nu}), A_\nu(n)] + \frac{a^2}{2} [A_\nu(n + \hat{\mu}), A_\mu(n + \hat{\nu})] + \frac{a^2}{2} [A_\mu(n), A_\nu(n)] \\
& \left. + \frac{a^2}{2} [A_\mu(n), A_\mu(n + \hat{\nu})] + \frac{a^2}{2} [A_\nu(n + \hat{\mu}), A_\nu(n)] \right) \quad (2.30)
\end{aligned}$$

Above, we have used the Baker-Campbell-Hausdorff formula to combine exponentials of non-commuting variables and ignored terms of order a^3 and higher. We can further reduce this expression by performing a Taylor series for the shifted fields and neglecting terms beyond a single power of a .

$$A_\mu(n + \hat{\nu}) \approx A_\mu(n) + a\partial_\nu A_\mu(n) + \dots \quad (2.31)$$

Taylor expanding these shifted fields will cause many terms in (2.30) to cancel, ultimately leaving

$$U_{\mu\nu} = \exp \left(ia^2\partial_\mu A_\nu(n) - ia^2\partial_\nu A_\mu(n) - a^2 [A_\mu(n), A_\nu(n)] \right). \quad (2.32)$$

Substituting (2.13), the continuum field strength tensor gives

$$U_{\mu\nu} = \exp \left(ia^2 F_{\mu\nu}(n) \right). \quad (2.33)$$

Finally, plugging this into (2.29) yields

$$S_G[U] = \frac{a^4}{2g^2} \sum_{n \in \Lambda} \sum_{\mu, \nu} \text{Tr} [F_{\mu\nu}(n)^2 + \dots]. \quad (2.34)$$

In the limit of zero lattice spacing, the extra terms in (2.34) vanish and the double sum becomes the usual space-time integral seen in the definition of the action from the QCD Lagrangian density. This concludes our discussion of the Wilson gauge action.

2.3 Fermions on the lattice

We will now discuss the addition of fermions to our lattice action. This examination will begin with the introduction and gauge-invariance check of the naive fermion action. The naive lattice Dirac operator will be presented and studied, naturally leading to the doubling problem. Conclusively, the Wilson term will be introduced to cure the doublers.

2.3.1 Naive lattice fermions

We have already introduced gauge links as the gluonic degrees of freedom on the lattice, now we turn our attention to fermions. We introduce lattice spinors on each of the lattice sites with Dirac, color, and flavor indices.

$$\psi^f(n)_{\alpha,c}, \bar{\psi}^f(n)_{\alpha,c} \quad (2.35)$$

Remember on the lattice, with the Euclidean metric, ψ and $\bar{\psi}$ are not directly related and must be treated as independent degrees of freedom. In following discussions we will suppress the indices on ψ for notational convenience. From (2.7), under a local color rotation, the lattice fermion fields must satisfy

$$\begin{aligned}\psi'(n) &= \Omega(n)\psi(n) \\ \bar{\psi}' &= \bar{\psi}(n)\Omega(n)^\dagger.\end{aligned}\tag{2.36}$$

Let us begin constructing a lattice fermion action by examining the free field fermion action in continuum QCD.

$$S_{F,free}[\psi, \bar{\psi}] = \int d^4x \bar{\psi}(x) (\gamma_\mu \partial_\mu + m) \psi(x),\tag{2.37}$$

The Euclidean metric and Euclidean γ matrices have been used above. In order to port this action to the lattice, the partial derivative is represented as a symmetric finite difference of fields at neighboring lattice points.

$$\partial_\mu \psi(x) \rightarrow \frac{1}{2a} [\psi(n + \hat{\mu}) - \psi(n - \hat{\mu})]\tag{2.38}$$

This allows the lattice free field fermion action to be written.

$$S_{F,free}[\psi, \bar{\psi}] = a^4 \sum_{n \in \Lambda} \bar{\psi}(n) \left[\sum_{\mu=1}^4 \gamma_\mu \frac{\psi(n + \hat{\mu}) - \psi(n - \hat{\mu})}{2a} + m\psi(n) \right]\tag{2.39}$$

Just as the continuum free field action, the mass term is gauge-invariant (seen trivially) and the derivative term is not.

$$\bar{\psi}'(n)\psi'(n + \hat{\mu}) = \bar{\psi}(n)\Omega(n)^\dagger\Omega(n + \hat{\mu})\psi(n + \hat{\mu})\tag{2.40}$$

However, similar to the introduction of the gauge-covariant derivative in the continuum, this term can be made gauge-invariant with a coupling to the gluon field.

$$\begin{aligned}
\bar{\psi}'(n)U'_\mu(n)\psi'(n+\hat{\mu}) &= \bar{\psi}(n)\Omega(n)^\dagger U'_\mu(n)\Omega(n+\hat{\mu})\psi(n+\hat{\mu}) \\
&= \bar{\psi}(n)U_\mu(n)\psi(n+\hat{\mu})
\end{aligned} \tag{2.41}$$

Here, we see the reason for postulating the link variable transformations as (2.19). Now we can write down a gauge-invariant fermion action, known as the naive fermion action.

$$S_F[\psi, \bar{\psi}, U] = a^4 \sum_{n \in \Lambda} \bar{\psi}(n) \left[\sum_{\mu=1}^4 \gamma_\mu \frac{U_\mu(n)\psi(n+\hat{\mu}) - U_{-\mu}(n)\psi(n-\hat{\mu})}{2a} + m\psi(n) \right] \tag{2.42}$$

Note that this is no longer in free field theory due to external gauge field (U) dependence.

2.3.2 The Dirac operator and the doubling problem

Let us focus now on a single quark flavor. Equation 2.42 may be rewritten in the form

$$S_F[\psi, \bar{\psi}, U] = a^4 \sum_{n, m \in \Lambda} \sum_{a, b, \alpha, \beta} \bar{\psi}(n, \alpha, a) D(n, \alpha, a | m, \beta, b) \psi(m, \beta, b) \tag{2.43}$$

$$D(n, \alpha, a | m, \beta, b) = \sum_{\mu=1}^4 (\gamma_\mu)_{\alpha\beta} \frac{U_\mu(n)_{ab} \delta_{n+\hat{\mu}, m} - U_{-\mu}(n)_{ab} \delta_{n-\hat{\mu}, m}}{2a} + m \delta_{\alpha\beta} \delta_{ab} \delta_{nm}. \tag{2.44}$$

$D(n, \alpha, a | m, \beta, b)$ is called the lattice Dirac operator, which may be viewed as a sparse matrix of dimension $12 \times N_L^3 \times N_T$. The only non-zero elements of the Dirac matrix are nearest neighbor connections on the lattice, this feature is called ultralocality. The fact

that the lattice Dirac operator is ultralocal naturally leads to the usefulness of a variety of sparse matrix techniques. Such numerical algorithms will be discussed in Chapter 4.

Now we analyze the Dirac matrix in free field theory ($U_\mu(n) = I$) and study its inverse. It is convenient to do this in momentum space so we apply a Fourier transformation to the space-time indices n and m .

$$\tilde{D}(p|q) = \frac{1}{N_L^3 N_T} \sum_{n,m \in \Lambda} e^{-iap \cdot n} D(n|m) e^{iaq \cdot m} \quad (2.45)$$

Above we have suppressed Dirac indices in favor of matrix/vector notation, note that since we are in free field, the color indices may be dropped. It is intuitive to write down the Fourier transformation on the lattice from the continuum expression, converting an integral to a sum, here we have simply employed it and leave details of its derivation to the reader in [31]. The Dirac matrix is only non-zero for forward and backward nearest neighbor connections, therefore (2.45) can be written as

$$\tilde{D}(p|q) = \frac{1}{N_L^3 N_T} \sum_{n \in \Lambda} e^{-ia(p-q) \cdot n} \left(\sum_{\mu=1}^4 \gamma_\mu \frac{e^{iaq_\mu} - e^{-iaq_\mu}}{2a} + mI \right) \quad (2.46)$$

$$= \delta(p - q) \tilde{D}(p),$$

$$\tilde{D}(p) = mI + \frac{i}{a} \sum_{\mu=1}^4 \gamma_\mu \sin(ap_\mu). \quad (2.47)$$

From the Delta function, the transformed Dirac operator is diagonal in momentum-space, mI is the 4×4 identity, scaled by the quark mass. The inverse Dirac matrix can be computed easily by inverting the Dirac structure for each momenta individually. In order to do so, we note that

$$\left(AI + i \sum_{\mu=1}^4 \gamma_{\mu} B_{\mu} \right)^{-1} = \frac{AI - i \sum_{\mu=1}^4 \gamma_{\mu} B_{\mu}}{A^2 + \sum_{\mu=1}^4 (B_{\mu})^2}, \quad (2.48)$$

where A is a scalar and B_{μ} is a Euclidean four-vector. This can be proved with a few lines of algebra.

$$\begin{aligned} \left(AI + i \sum_{\mu=1}^4 \gamma_{\mu} B_{\mu} \right)^{-1} \left(AI + i \sum_{\mu=1}^4 \gamma_{\mu} B_{\mu} \right) &= \frac{AI - i \sum_{\mu=1}^4 \gamma_{\mu} B_{\mu}}{A^2 + \sum_{\mu=1}^4 (B_{\mu})^2} \left(AI + i \sum_{\mu=1}^4 \gamma_{\mu} B_{\mu} \right) \\ I &= \frac{A^2 I + \sum_{\mu=1}^4 (B_{\mu})^2 I}{A^2 + \sum_{\mu=1}^4 (B_{\mu})^2} \\ I &= I \end{aligned}$$

In the middle step we have used the Euclidean gamma anti-commutation relations A.2. Equation 2.48 can be applied directly to compute the Dirac matrix inverse.

$$\tilde{D}(p)^{-1} = \frac{mI - \frac{i}{a} \sum_{\mu=1}^4 \gamma_{\mu} \sin(ap_{\mu})}{m^2 + a^{-2} \sum_{\mu=1}^4 \sin(ap_{\mu})^2} \quad (2.49)$$

Akin to continuum quantum field theory, this is the momentum-space quark propagator. We should check that in the right limits, this propagator resembles the continuum fermion propagator $\frac{i(\not{p}+m)}{p^2-m^2+i\epsilon}$. To simplify this analysis, we use the chiral mass ($m = 0$), followed by the zero lattice-spacing limit.

$$\tilde{D}(p)^{-1}|_{m=0, a \rightarrow 0} = \frac{-i \sum_{\mu=1}^4 \gamma_{\mu} p_{\mu}}{p^2} \quad (2.50)$$

This agrees perfectly with the continuum expression up to a sign, due to the difference in metrics. It is clear that in these limits, the propagator has a pole at $p = (0, 0, 0, 0)$, matching once again the continuum behavior. Unfortunately, with massless fermions, (2.49) also has poles at $p = (0, 0, 0, \frac{\pi}{a}), (0, 0, \frac{\pi}{a}, 0), \dots, (0, \frac{\pi}{a}, \frac{\pi}{a}, \frac{\pi}{a}), (\frac{\pi}{a}, \frac{\pi}{a}, \frac{\pi}{a}, \frac{\pi}{a})$, resulting in 16

total poles. The 15 excess poles, known as fermion doublers, do not correspond to physical particles and pose a serious issue for lattice simulations.

2.3.3 Wilson fermions

One possible solution to remove the doublers and leave the physical pole at $p_\mu = 0$ intact, is to add an extra term to the Dirac matrix.

$$\tilde{D}(p) = mI + \frac{i}{a} \sum_{\mu=1}^4 \gamma_\mu \sin(ap_\mu) + a^{-1} \sum_{\mu=1}^4 [1 - \cos(ap_\mu)] I \quad (2.51)$$

This additional term is known as the Wilson term. It does not affect the pole at $p_\mu = 0$ in any way, however the cosine does add $\frac{2}{a}$ contributions to all the other poles that arise from the sine. The mass of the physical quark remains m , while all the doublers acquire an additional mass:

$$m + \frac{2l}{a}, \quad (2.52)$$

where l is an integer from 1 to 4 that totals up the contributions from $p_\mu = \frac{\pi}{a}$. In the limit of zero lattice-spacing, the doublers become infinitely heavy, decoupling from the theory. Additionally, repeating the examination of the momentum-space propagator in free field theory gives only one pole at $p_\mu = 0$. Converting back to position-space, we may read off $1 - \cos(ap_\mu)$ by comparing to the Fourier transformation with the original $\sin(ap_\mu)$ term.

$$\sum_{\mu=1}^4 \frac{2\delta_{ab}\delta_{nm} - U_\mu(n)_{ab}\delta_{n+\hat{\mu},m} - U_{-\mu}(n)_{ab}\delta_{n-\hat{\mu},m}}{2a} \quad (2.53)$$

Combining the Wilson term with the naive Dirac operator, we may now write down the full expression for Wilson fermions.

$$\begin{aligned}
D(n, \alpha, a|m, \beta, b) &= \sum_{\mu=1}^4 (\gamma_{\mu})_{\alpha\beta} \frac{U_{\mu}(n)_{ab} \delta_{n+\hat{\mu},m} - U_{-\mu}(n)_{ab} \delta_{n-\hat{\mu},m}}{2a} + m \delta_{\alpha\beta} \delta_{ab} \delta_{nm} \\
&+ \sum_{\mu=1}^4 \frac{2\delta_{ab} \delta_{nm} - U_{\mu}(n)_{ab} \delta_{n+\hat{\mu},m} - U_{-\mu}(n)_{ab} \delta_{n-\hat{\mu},m}}{2a} \\
&= \left(m + \frac{4}{a}\right) \delta_{\alpha\beta} \delta_{ab} \delta_{nm} - \frac{1}{2a} \sum_{\mu=\pm 1}^{\pm 4} (I - \gamma_{\mu})_{\alpha\beta} U_{\mu}(n)_{ab} \delta_{n+\hat{\mu},m} \quad (2.54)
\end{aligned}$$

Above we have defined $\gamma_{-\mu} = -\gamma_{\mu}$. With this we have achieved a fully working, gauge-invariant lattice action for fermions. Additionally, we have demonstrated that in the continuum limit, we recover the correct fermion propagator, free of doublers. There are many other kinds of lattice fermions such as Staggered/HISQ [33], Domain Wall [34, 35, 36], Twisted-mass [37], and Clover [38, 39]. We encourage the reader to read the standard literature for a complete and rigorous treatment of fermions, as we will finish our discussion by briefly examining only Clover fermions.

2.4 Symanzik improvement

When formulating the fermion lattice action in the previous section, one of the objects we introduced was the symmetric discretized derivative, in (2.38). In doing so however, we have introduced discretization effects, or errors from a finite-lattice spacing. The same statement is also true in our treatment of proving that the summed trace of plaquettes approaches the continuum QCD gluon action. Our fermion action has $\mathcal{O}(a)$ discretization effects, while the gluon action contains $\mathcal{O}(a^2)$ effects. Although these effects disappear in the continuum limit, in practice a computer simulation must always work in finite a , followed by an extrapolation across different numerical simulations to achieve the continuum limit. Naturally, the higher order in which discretization errors are present, the quicker

these “lattice artifacts”, or unwanted discretization terms tend to zero.

It should be noted that our choice of lattice action is not unique, and that other actions with the same continuum behavior work just as well, some with lower discretization errors. For example, an additional term can be added to the Wilson fermion action and with precise matching of its coefficients, its discretization errors may come in with $\mathcal{O}(a^2)$ instead of $\mathcal{O}(a)$. A systematic application of this method is called Symanzik improvement. In order to realize full $\mathcal{O}(a)$ improvement, the discretization errors of lattice correlators (which will be introduced in the next chapter) must also be treated. Such treatise is beyond the scope of this section.

2.4.1 Toy model

We will start by improving a simple discretized toy model to attain a basic prescription of the improvement process. Many steps below will be identical to applying Symanzik improvement to Lattice QCD and will not be redone in the following subsection. Let us introduce a symmetric and discretized finite difference, approximating the derivative of a single variable function in one dimension.

$$\frac{f(x+a) - f(x-a)}{2a} = f'(x) + aC^{(1)}(x) + a^2C^{(2)}(x) + a^3C^{(3)}(x) + a^4C^{(4)}(x) + \mathcal{O}(a^5) \quad (2.55)$$

Assuming f is dimensionless, the coefficients $C^{(1)}(x), C^{(2)}(x), C^{(3)}(x), C^{(4)}(x), \dots$ have units of length: $[l]^{-2}, [l]^{-3}, [l]^{-4}, [l]^{-5} \dots$. We may use a Taylor series to match the derivatives of f with the coefficients on the right side of (2.55).

$$f(x \pm a) = f(x) \pm af'(x) + \frac{a^2}{2}f''(x) \pm \frac{a^3}{6}f'''(x) + \dots \quad (2.56)$$

All even powers of a will cancel in the expansion of the numerator of (2.55). Consequentially, only odd derivatives of f will come in on the right side, matched with the coefficients of the even powers of a .

$$C^{(1)}(x), C^{(3)}(x), \dots = 0 \quad (2.57)$$

Accounting for the factor of $2a$ in the denominator, we find

$$C^{(3)}(x) = \frac{1}{6}f'''(x). \quad (2.58)$$

The key to the improvement process is to add to the left side of (2.55), a discretized term that cancels the correction on the right side to a given order. Therefore, for $\mathcal{O}(a^2)$ improvement we add

$$\frac{f(x+a) - f(x-a)}{2a} + a^2 D E^{(3)}[f](x) = f'(x) + \mathcal{O}(a^4), \quad (2.59)$$

where D is a constant that needs to be determined and $E^{(3)}[f](x) \approx f'''(x) + \mathcal{O}(a^2)$.

Using (2.56) again, one possible candidate function could be

$$D = -\frac{1}{6}, \quad E^{(3)}[f](x) = \frac{f(x+2a) + 2f(x-a) - f(x-2a) - 2f(x+a)}{2a^3}. \quad (2.60)$$

Note that this choice is not unique, however it is one possible discretized term that yields $\mathcal{O}(a^2)$ improvement. The same exact procedure is followed in Lattice QCD, with one subtlety: determining the coefficient D . In our toy model, D was easily identified with a few lines of algebra, however due to the gluon dressing of the quark propagators, discovering the analogous coefficients in QCD is much more complicated. Typically these coefficients are found either perturbatively or with a non-perturbative improvement procedure.

2.4.2 Implementation to Lattice QCD

Now we will apply this set of ideas to Lattice QCD. Our full QCD action involves the Wilson gauge action in (2.29) and the Wilson Dirac matrix in (2.54). As stated previously, the fermion part contains $\mathcal{O}(a)$ discretization errors and the gluon part has $\mathcal{O}(a^2)$ errors.

Repeating the steps in the preceding subsection, we write down correction terms to our continuum action. These terms are ordered with respect to their dimension and should satisfy all the symmetries of the starting QCD action. This is analogous to adding higher order operators in effective field theory.

$$S_{effective} = \int d^4x (L^{(0)}(x) + aL^{(1)}(x) + a^2L^{(2)}(x) + \dots) \quad (2.61)$$

Above, $L^{(0)}$ is the QCD Lagrangian defined by (2.3) and (2.15). The terms $L^{(1)}$ and onward are the corrections, built out of quark and gluon fields so that their dimension matches $[L^{(k)}] = [L]^{-(4+k)}$. Imposing the symmetries of the QCD lattice action, it may be shown that the dimension-5 correction, $L^{(1)}$, can be written as linear combinations of the following:

$$\begin{aligned} L_1^{(1)} &= \bar{\psi}(x)\sigma_{\mu\nu}F_{\mu\nu}\psi(x) \\ L_2^{(1)} &= \bar{\psi}(x)\vec{D}_\mu(x)\vec{D}_\mu(x)\psi(x) + \bar{\psi}(x)\overleftarrow{D}_\mu(x)\overleftarrow{D}_\mu(x)\psi(x) \\ L_3^{(1)} &= \bar{\psi}(x)m\text{Tr}[F_{\mu\nu}F_{\mu\nu}]\psi(x) \\ L_4^{(1)} &= m\bar{\psi}(x)\gamma_\mu\vec{D}_\mu(x)\psi(x) - m\bar{\psi}(x)\gamma_\mu\overleftarrow{D}_\mu(x)\psi(x) \\ L_5^{(1)} &= m^2\bar{\psi}(x)\psi(x). \end{aligned} \quad (2.62)$$

Here $\sigma_{\mu\nu} = \frac{1}{2i}[\gamma_\mu, \gamma_\nu]$, and we have also introduced non-symmetric left and right derivatives. From the field equation of motion $(\gamma_\mu D_\mu + m)\psi = 0$, further constraints can

be placed on linear combinations of these operators.

$$L_1^{(1)} - L_2^{(1)} + 2L_5^{(5)} = 0 \quad (2.63)$$

$$L_4^{(1)} + 2L_5^{(5)} = 0 \quad (2.64)$$

This allows $L_2^{(1)}$ and $L_4^{(1)}$ to be eliminated [40], resulting in only $L_1^{(1)}$, $L_3^{(1)}$, and $L_5^{(1)}$. Notice that using the equations of motion to eliminate these terms only gives $\mathcal{O}(a)$ improvement for on-shell quantities. Furthermore, $L_3^{(1)}$ and $L_5^{(1)}$ already appear in the original action, a simple redefinition of the bare parameters m and β can be done to represent the addition of these terms. The only remaining term is $L_1^{(1)}$, which we add to the lattice action.

$$S = S_{Wilson} + c_{sw} a^5 \sum_{n \in \Lambda} \sum_{\mu < \nu} \bar{\psi}(n) \frac{1}{2} \sigma_{\mu\nu} \hat{F}_{\mu\nu} \psi(n) \quad (2.65)$$

The coefficient c_{sw} is called the Sheikholeslami-Wohlert coefficient [39] and $\hat{F}_{\mu\nu}$ is a discretized version of the gluon field strength tensor. The Sheikholeslami-Wohlert coefficient may be determined through a perturbative matching procedure or non-perturbatively with the partially conserved axial current relation (PCAC) on the Schrödinger functional [41]. A particularly useful discretization of $F_{\mu\nu}$ is

$$\hat{F}_{\mu\nu} = -\frac{i}{8a^2} [Q_{\mu\nu}(n) - Q_{\nu\mu}(n)] \quad (2.66)$$

$$Q_{\mu\nu}(n) = U_{\mu\nu}(n) + U_{\nu,-\mu}(n) + U_{-\mu,-\nu}(n) + U_{-\nu,\mu}(n). \quad (2.67)$$

The shape of $Q_{\mu\nu}$ resembles a four-leaf clover, as such the discretized version of $L_1^{(1)}$

is known as the clover term.

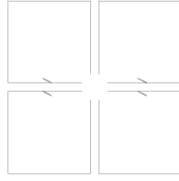


FIG. 2.2: A pictorial representation of $Q_{\mu\nu}$ in the $\mu - \nu$ plane.

We have now achieved $\mathcal{O}(a)$ improvement for the Wilson lattice action. We remind the reader that in order for calculations to be fully $\mathcal{O}(a)$ improved, the Euclidean correlators we compute on the lattice must also be improved. We will not discuss this improvement, however the next chapter will go in depth regarding the measurement of observables, such as hadron matrix elements on the lattice.

CHAPTER 3

Computing Matrix Elements in Lattice QCD

As mentioned in Chapter 1, hadron structure observables are extracted from Euclidean correlators on the lattice. We discuss how to construct two-point and three-point correlators using the nucleon as an example, although these methods are useful in studying a variety of different hadrons. We start by introducing the interpolating fields and creation/annihilation operators for the nucleon. Other details of two-point function generation such as contractions and smearing follow. Subsequently, we introduce the reader to three-point functions and categorize the possible Wick contractions into two classifications: connected and disconnected diagrams. This chapter is then concluded with a discussion of excited state contamination and techniques to isolate the ground state matrix element based on correlated fits.

3.1 Two-point functions

The basic ingredient for hadron spectroscopy calculations is the two-point correlator. This correlation function gives access to the mass of the corresponding hadron. One of the impressive successes of Lattice QCD over the past few decades has been in producing the hadron mass spectrum that closely matches experimental results. In the following subsections we outline the basic steps involved in constructing Euclidean two-point correlators on the lattice.

3.1.1 Nucleon interpolating field

The proton and neutron are $J = \frac{1}{2}$, $P = +1$, $I = \frac{1}{2}$ particles, with $I_z = +\frac{1}{2}, -\frac{1}{2}$ respectively. Their masses are nearly identical at approximately 940 MeV. The proton has one fundamental unit of electric charge, while the neutron is electrically neutral, as its namesake implies. The quark flavor content of the proton is uud , while the neutron's is ddu , consistent with their net charges. Due to approximate isospin symmetry the proton/neutron masses are very close. There is a small isospin symmetry breaking contribution since the up and down quarks do not have identically degenerate masses. Additionally, electromagnetic interactions contribute differently to the proton's and neutron's mass. However, these effects are subleading compared to the strong interactions, as such we do not differentiate between the proton or neutron and merely categorize them together as nucleons.

There are three linearly independent local interpolating fields that are consistent with the quantum numbers of these nucleons. We give them the subscripts 1, 2, 3 and write them as

$$\mathcal{O}_{p,1}(n) = \epsilon_{abc} u(n)_a [u(n)_b^T C \gamma_5 d(n)_c] \quad (3.1)$$

$$\mathcal{O}_{p,2}(n) = \epsilon_{abc} \gamma_5 u(n)_a [u(n)_b^T C d(n)_c] \quad (3.2)$$

$$\mathcal{O}_{p,3}(n) = i \epsilon_{abc} u(n)_a [u(n)_b^T \gamma_4 C \gamma_5 d(n)_c] \quad (3.3)$$

$$\mathcal{O}_{n,1}(n) = \epsilon_{abc} d(n)_a [d(n)_b^T C \gamma_5 u(n)_c] \quad (3.4)$$

$$\mathcal{O}_{n,2}(n) = \epsilon_{abc} \gamma_5 d(n)_a [d(n)_b^T C u(n)_c] \quad (3.5)$$

$$\mathcal{O}_{n,3}(n) = i \epsilon_{abc} d(n)_a [d(n)_b^T \gamma_4 C \gamma_5 u(n)_c], \quad (3.6)$$

where a, b, c , are color indices and C is the charge conjugation matrix, represented as $C = i\gamma_2\gamma_4$ in the Weyl (chiral) basis. The Dirac indices above have been suppressed in favor of matrix/vector notation. The transposition, T , converts a column four-vector Dirac spinor into a row one. We will now limit our discussion to the first interpolating field from (3.1), although there is nothing unique about this field and the contractions in the forthcoming subsection are applicable to any of the three interpolators (or linear combinations of them). These interpolators are not quite the nucleon creation operators since information about parity is missing. In the case of forward propagation, we may project to definite parity $P = +1$ by using the proton creation operator

$$\bar{\mathcal{O}}_p(n) = \epsilon_{abc} [\bar{u}(n)_a C \gamma_5 \bar{d}(n)_b^T] \bar{u}(n)_c (I + \gamma_4). \quad (3.7)$$

The neutron creation operator is identical to the above with the up and down quark spinors switched. The forward proton annihilation operator is similarly

$$\mathcal{O}_p(n) = \epsilon_{abc} (I + \gamma_4) u(n)_a [u(n)_b^T C \gamma_5 d(n)_c]. \quad (3.8)$$

It is important to note that these operators couple to all the particles with the same quantum numbers as the nucleon. It is in general, a highly non-trivial task to determine the best operators that yield the highest overlap with matrix elements of interest in a calculation. There has been an ongoing evolution in the Lattice community for decades in finding the most appropriate operators or techniques and we do not focus on those details in this work. For now, we continue with our basic choice of the nucleon creation/annihilation operators and use them to construct the proton two-point correlation function.

3.1.2 Contractions

Now the ingredients are in place to be able to compute the proton propagator. We perform the Wick contractions and write the proton two-point function in terms of the quark propagators. We remind the reader that, as shown in Chapter 2, the quark propagator is simply the inverse of the Dirac matrix.

$$\begin{aligned}
& \frac{1}{Z_F[U]} \int \mathcal{D}[\psi, \bar{\psi}] e^{-S_F[\psi, \bar{\psi}, U]} \mathcal{O}_p(n)_\alpha \bar{\mathcal{O}}_p(m)_\alpha \\
= & -\frac{1}{Z_F[U]} \int \mathcal{D}[\psi, \bar{\psi}] e^{-S_F[\psi, \bar{\psi}, U]} \bar{\mathcal{O}}_p(m)_\alpha \mathcal{O}_p(n)_\alpha \\
= & -\frac{1}{Z_F[U]} \int \mathcal{D}[\psi, \bar{\psi}] e^{-S_F[\psi, \bar{\psi}, U]} \epsilon_{abc} \epsilon_{a'b'c'} [\bar{u}(n)_a C \gamma_5 \bar{d}(n)_b^T] \bar{u}(n)_c (I + \gamma_4) u(n)_{c'} [u(n)_a^T C \gamma_5 d(n)_{b'}] \\
= & \epsilon_{abc} \epsilon_{a'b'c'} (C \gamma_5)_{\alpha' \beta'} (C \gamma_5)_{\alpha \beta} (I + \gamma_4)_{\gamma \gamma'} D_d^{-1}(n, \beta', b' | m, \beta, b) \times \\
& [D_u^{-1}(n, \alpha', a' | m, \alpha, a) D_u^{-1}(n, \gamma', c' | m, \gamma, c) - D_u^{-1}(n, \alpha', a' | m, \gamma, c) D_u^{-1}(n, \gamma', c' | m, \alpha, a)] \quad (3.9)
\end{aligned}$$

Where

$$Z_F[U] = \int \mathcal{D}[\psi, \bar{\psi}] e^{-S_F[\psi, \bar{\psi}, U]}, \quad (3.10)$$

also known as the fermion determinant. Details of the derivation of (3.10) can be found in Appendix B. We have explicitly written the spinor indices for the creation/annihilation operators in the first line of (3.9), where the repeated α is summed over, additionally in the third line $(I + \gamma_4)^2 = (I + \gamma_4)$ was used. Notice the minus sign between the two terms, arising from anticommuting the quark fields.

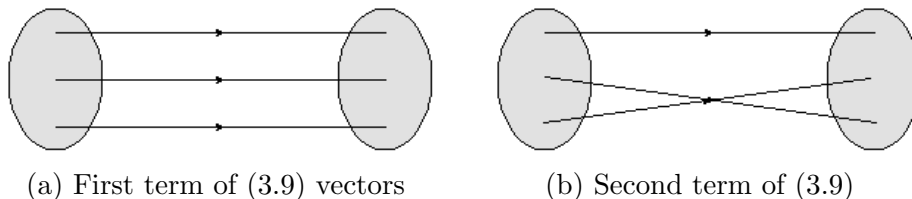


FIG. 3.1: Here we schematically show the contractions of the nucleon two-point correlator. The top fermion line is the down quark, the bottom two are up quarks.

3.1.3 Quark sources and smearing

Equation 3.9 is the recipe for creating a proton two-point function on the lattice, however it involves inverting the Dirac operator. Typical volumes for Lattice simulations are currently $32^3 \times 64$, $48^3 \times 96$, or larger, resulting in matrices of dimension 25,000,000², 127,000,000², or higher. Although the Dirac matrix is sparse, its inverse is in general not, resulting in an incredibly large and dense matrix. It would take an unrealistic amount of time to compute this matrix for even one gauge field configuration, of which many are needed to obtain a full result. One possible solution to this issue is to use point-sources.

The point-source method involves only computing the quark propagator from a fixed lattice site, Dirac index, and color.

$$D^{-1}(n, \beta, b | m_i, \alpha_i, a_i) = \sum_{m, \alpha, a} D^{-1}(n, \beta, b | m, \alpha, a) S_{m_i, \alpha_i, a_i}(m, \alpha, a) \quad (3.11)$$

$$S_{m_i, \alpha_i, a_i}(m, \alpha, a) = \delta(m - m_i) \delta_{\alpha \alpha_i} \delta_{a a_i} \quad (3.12)$$

Inverting 12 of these point-sources so that each spin/color combination for a specified lattice site is included yields the point-to-all propagator. This step may be repeated for various lattice sites to acquire reasonable statistics for the desired observable without the need to compute the full inverse of an extremely large matrix. Put simply, this method relies on only finding a few columns of D^{-1} , rather than the full object. Nevertheless, even with this method, inversions can often be the bottleneck in Lattice QCD calculations. This fact has caused a wealth of algorithms to emerge, aimed at decreasing this solver time. We will study such algorithms in detail in the following chapter.

As mentioned earlier, the proton interpolator will contain all physical states with the allowed symmetries. It is expected that quarks are distributed smoothly over the ground state proton, therefore a simple technique that helps optimize the interpolator is quark smearing. This is achieved by applying a smearing kernel to a point-source vector before inversion (source smearing), then inserting it again when the proton propagator is constructed out of the quark fields (sink smearing). Smearing may be viewed as a way of suppressing the interpolating field's coupling with particles in motion with the same quantum numbers as the proton, thereby reducing resolution of momentum in exchange for higher accuracy in position.

3.2 Three-point functions

In this section we introduce and work out the contractions for the proton three-point function. This is the lattice quantity relevant for extracting matrix elements of a current interaction with a proton. This current insertion J_μ is coupled to the proton through quark bilinears, therefore there are separate contributions from up, down, and strange.

$$\bar{u}J_\mu u \tag{3.13}$$

$$\bar{d}J_\mu d \tag{3.14}$$

$$\bar{s}J_\mu s \tag{3.15}$$

The insertion happens at a time between the creation (source) and annihilation (sink) of the proton propagator. Thereafter writing the contractions, the resulting diagrams are categorized into two types, connected and disconnected. We finish this section by noting a key difference between connected and disconnected contributions, which causes disconnected diagrams to be considerably more computationally expensive.

3.2.1 Up quark contractions

The up quark contractions for the three-point function are written below.

$$\begin{aligned} & \frac{1}{Z_F[U]} \int \mathcal{D}[\psi, \bar{\psi}] e^{-S_F[\psi, \bar{\psi}, U]} \mathcal{O}_p(m)_\alpha \bar{u}_{\delta' d'} J_\mu(o, \delta', d' | o, \delta, d) u_{\delta d} \bar{\mathcal{O}}_p(n)_\alpha \\ = & \epsilon_{abc} \epsilon_{a'b'c'} (C\gamma_5)_{\alpha'\beta'} (C\gamma_5)_{\alpha\beta} (I + \gamma_4)_{\gamma\gamma'} J_\mu(o, \delta', d' | o, \delta, d) D_d^{-1}(n, \beta', b' | m, \beta, b) \times \\ & [D_u^{-1}(o, \delta', d' | m, \gamma, c) D_u^{-1}(n, \gamma', c' | o, \delta, d) D_u^{-1}(n, \alpha', a' | m, \alpha, a) \\ & - D_u^{-1}(o, \delta', d' | m, \gamma, c) D_u^{-1}(n, \gamma', c' | m, \alpha, a) D_u^{-1}(n, \alpha', a' | o, \delta, d) \\ & - D_u^{-1}(o, \delta', d' | m, \alpha, a) D_u^{-1}(n, \gamma', c' | o, \delta, d) D_u^{-1}(n, \alpha', a' | m, \gamma, c) \\ & + D_u^{-1}(o, \delta', d' | m, \alpha, a) D_u^{-1}(n, \gamma', c' | m, \gamma, c) D_u^{-1}(n, \alpha', a' | o, \delta, d) \\ & - D_u^{-1}(o, \delta', d' | o, \delta, d) D_u^{-1}(n, \gamma', c' | m, \gamma, c) D_u^{-1}(n, \alpha', a' | m, \alpha, a) \\ & + D_u^{-1}(o, \delta', d' | o, \delta, d) D_u^{-1}(n, \gamma', c' | m, \alpha, a) D_u^{-1}(n, \alpha', a' | m, \gamma, c)] \end{aligned} \tag{3.16}$$

As can be seen from the last 2 lines of (3.16), one possible type of contraction allows for the \bar{u} and u from the operator insertion to connect with each other, we refer to these diagrams as “disconnected”. The propagator $D_u^{-1}(o, \delta', d' | o, \delta, d)$ loops a u quark from one space-time point back to the same space-time point. At the end of this section it will be shown that this feature of the diagram makes it exceptionally difficult to calculate. We note that these are not truly disconnected Feynman diagram; the quark loop is coupled to the proton propagator through the QCD vacuum, in other words the loop is “dressed” by all possible gluon lines that connect to the proton two-point function.

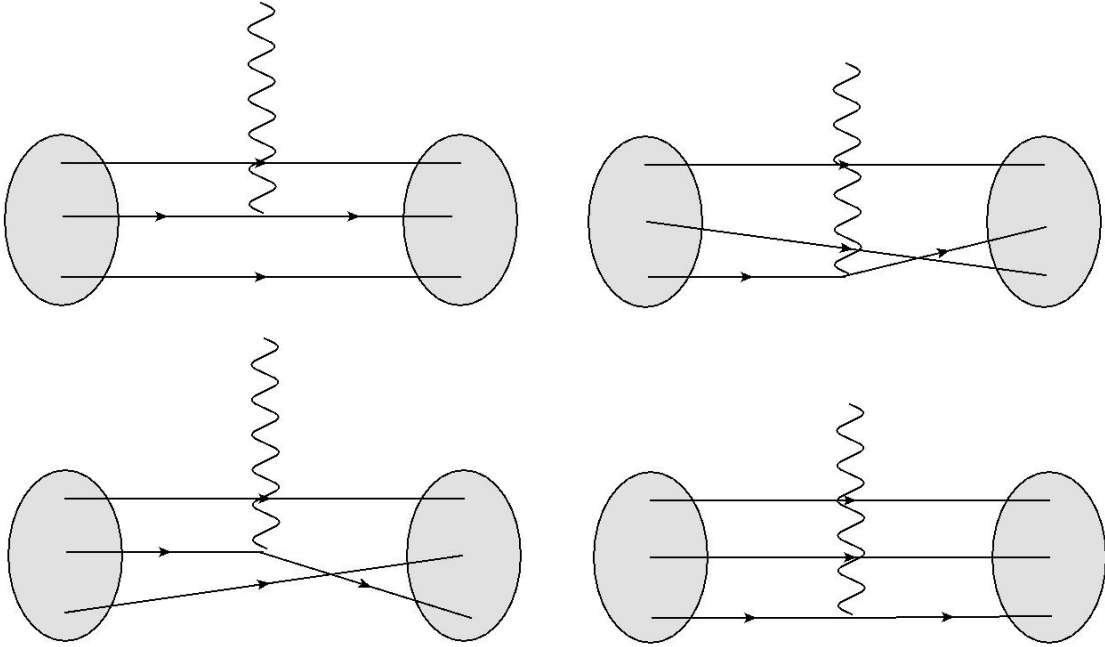


FIG. 3.2: The first four terms of the up quark contractions in (3.16) are represented pictorially above. These diagrams only illustrate the Wick contractions, as all possible gluons connect the fermion lines. The top fermion is the down quark and the bottom two are up quarks in the proton. These diagrams are typically classified as the connected pieces.

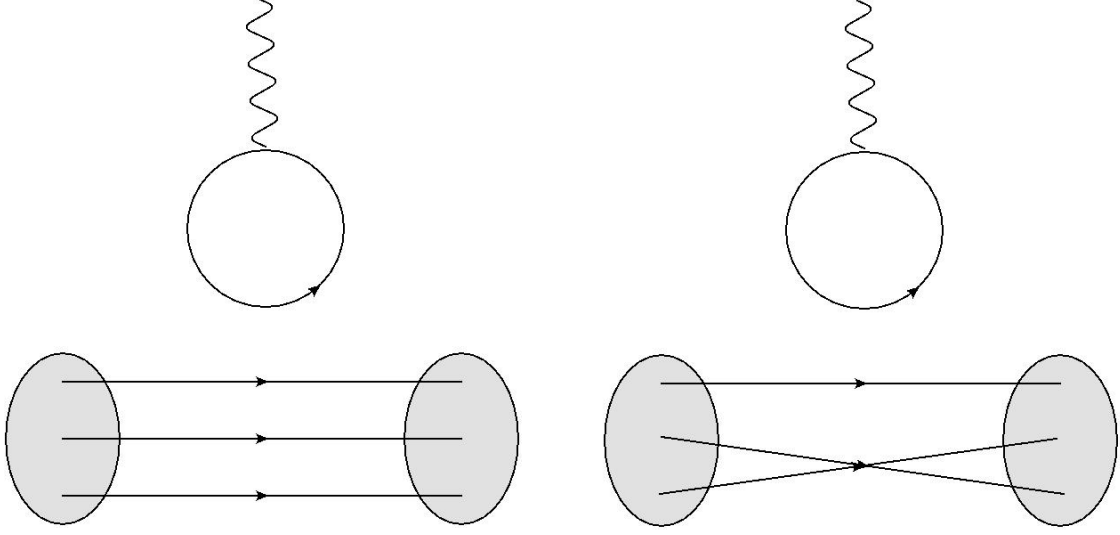


FIG. 3.3: The last two terms of (3.16) are shown schematically. The loop is comprised of an up quark propagating from around the same timeslice. These are categorized as the disconnected contributions of the up quark.

3.2.2 Down quark contractions

Next, we redo the same calculation as (3.2) for the down quark bilinear.

$$\begin{aligned}
& \frac{1}{Z_F[U]} \int \mathcal{D}[\psi, \bar{\psi}] e^{-S_F[\psi, \bar{\psi}, U]} \mathcal{O}_p(m)_\alpha \bar{d}_{\delta' d'} J_\mu(o, \delta', d' | o, \delta, d) d_{\delta d} \bar{\mathcal{O}}_p(n)_\alpha \\
= & \epsilon_{abc} \epsilon_{a'b'c'} (C\gamma_5)_{\alpha'\beta'} (C\gamma_5)_{\alpha\beta} (I + \gamma_4)_{\gamma\gamma'} J_\mu(o, \delta', d' | o, \delta, d) \times \\
& [D_d^{-1}(n, \beta', b | o, \delta, d) D_d^{-1}(o, \delta', d' | m, \beta, b) D_u^{-1}(n, \alpha', a' | m, \alpha, a) D_u^{-1}(n, \gamma', c' | m, \gamma, c) \\
- & D_d^{-1}(n, \beta', b | o, \delta, d) D_d^{-1}(o, \delta', d' | m, \beta, b) D_u^{-1}(n, \alpha', a' | m, \gamma, c) D_u^{-1}(n, \gamma', c' | m, \alpha, a) \\
- & D_d^{-1}(n, \beta', b' | m, \beta, b) D_d^{-1}(o, \delta', d' | o, \delta, d) D_u^{-1}(n, \alpha', a' | m, \alpha, a) D_u^{-1}(n, \gamma', c' | m, \gamma, c) \\
+ & D_d^{-1}(n, \beta', b' | m, \beta, b) D_d^{-1}(o, \delta', d' | o, \delta, d) D_u^{-1}(n, \alpha', a' | m, \gamma, c) D_u^{-1}(n, \gamma', c' | m, \alpha, a)]
\end{aligned} \tag{3.17}$$

Once again both connected and disconnected diagrams appear.

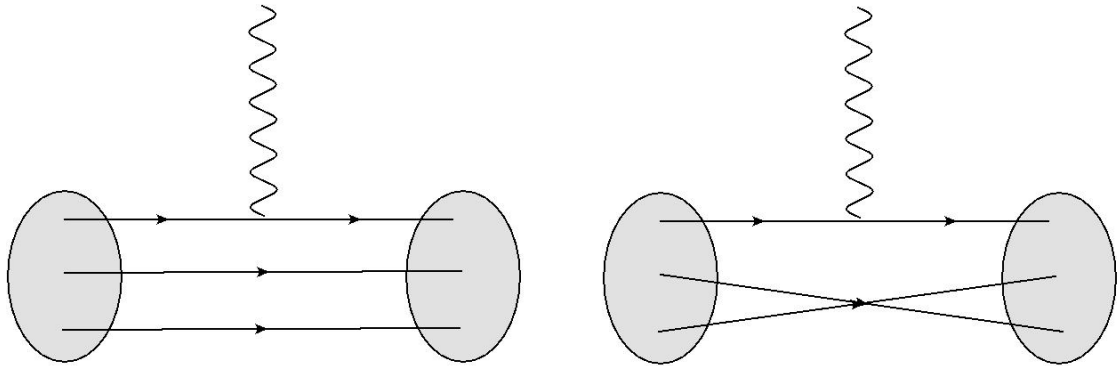


FIG. 3.4: The first two terms in (3.17) are shown. Since the proton has only one down quark, only two different contractions are present in the connected case.

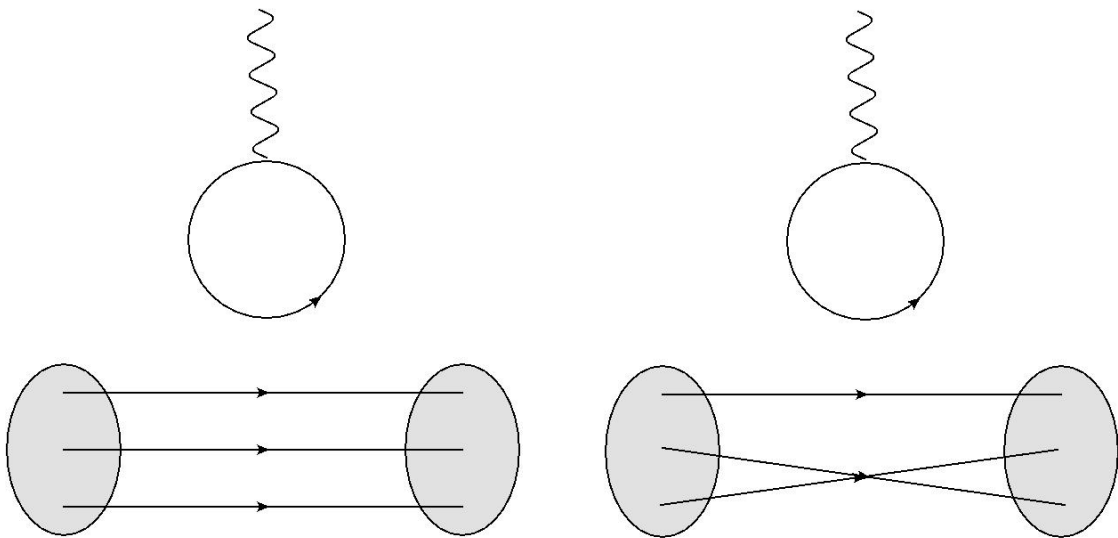


FIG. 3.5: Here we show the last two terms of (3.17). The loop is made of a propagating down quark. Due to isospin symmetry, usually only one computation is done for Figures 3.3 and 3.5.

3.2.3 Strange quark contractions

Finally, we write down the strange quark contractions. Since there are no strange quarks in the valence sector of the proton, only disconnected diagrams contribute to these matrix elements.

$$\begin{aligned}
& \frac{1}{Z_F[U]} \int \mathcal{D}[\psi, \bar{\psi}] e^{-S_F[\psi, \bar{\psi}, U]} \mathcal{O}_p(m)_\alpha \bar{s}_{\delta' d'} J_\mu(o, \delta', d' | o, \delta, d) s_{\delta d} \bar{\mathcal{O}}_p(n)_\alpha \\
= & \epsilon_{abc} \epsilon_{a'b'c'} (C\gamma_5)_{\alpha'\beta'} (C\gamma_5)_{\alpha\beta} (I + \gamma_4)_{\gamma\gamma'} J_\mu(o, \delta', d' | o, \delta, d) \times \\
& [D_s^{-1}(o, \delta', d' | o, \delta, d) D_d^{-1}(n, \beta', b' | m, \beta, b) D_u^{-1}(n, \alpha', a' | m, \alpha, a) D_u^{-1}(n, \gamma', c' | m, \gamma, c) \\
- & D_s^{-1}(o, \delta', d' | o, \delta, d) D_d^{-1}(n, \beta', b' | m, \beta, b) D_u^{-1}(n, \alpha', a' | m, \gamma, c) D_u^{-1}(n, \gamma', c' | m, \alpha, a)]
\end{aligned} \tag{3.18}$$

This makes disconnected diagrams central in addressing physics such as the strange content of the nucleon.

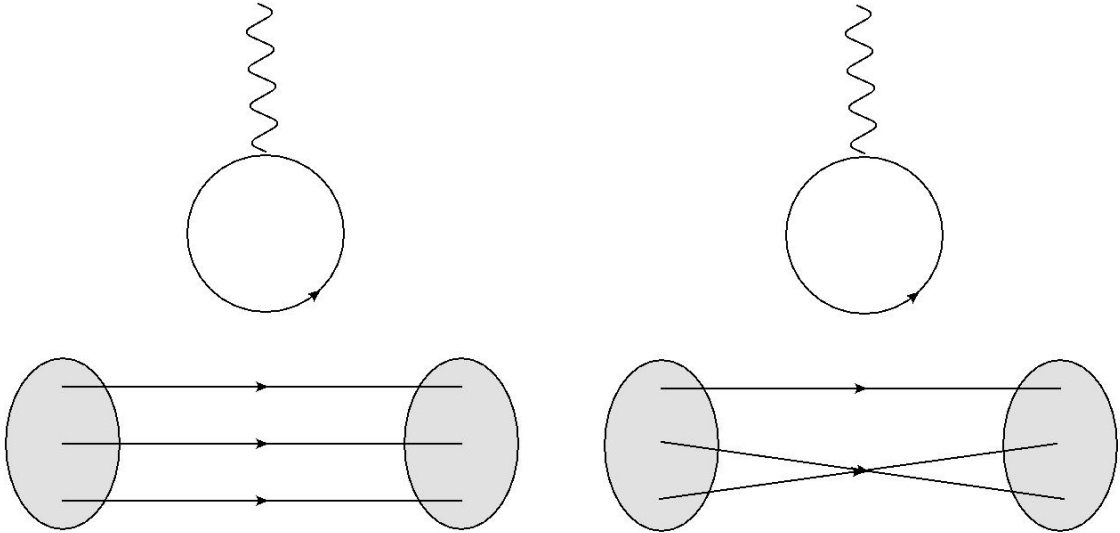


FIG. 3.6: The only types of contributions from strange quark contractions are disconnected.

3.2.4 Difference between connected and disconnected diagrams

It is clear by surveying the contraction plots of the previous subsection that the operator insertion in the connected diagrams contracts with different quark lines. This distinction is vital in understanding the disparity in computational cost between connected and disconnected contributions. Inverting a source vector, as stated earlier, gives the forward quark propagator or the point-to-all propagator. Additionally, γ_5 hermiticity can be used to attain the backward or all-to-point propagator. This allows for the “free” ends of forward and backward quark propagators to be tied with the operator insertion $J_\mu(o, \delta', d'|o, \delta, d)$ [42]. This can also be viewed as cutting one of the quarks in the proton two-point correlator to insert a current. After this step, a trace over the spin/color degrees of freedom yields the three-point correlator.

By contrast, the disconnected diagrams encompass a quark loop propagating back to the same space-time point. The operator insertion is contracted with both ends of the same quark line. Because the momentum flow is unspecified through this loop, a sum over all possible momenta is needed. Therefore, the previous trick of combining the current with the open ends of the forward and backward propagators cannot be applied. Instead, the trace of the full matrix-inverse is required, or the trace of the all-to-all propagator.

Computing the trace of the all-to-all propagator exactly, even for only a single time-slice, would take an enormous amount of computing power. The typical approach to calculating a quark loop is to instead employ a stochastic Monte Carlo trace estimator. While this reduces computational costs, convergence of this estimator can be an issue. Further details of this will be elucidated in the ensuing chapter.

3.3 Reducing excited state contamination

As mentioned in Section 3.1.1, the nucleon interpolating field we use contains contributions from not just the ground state nucleon, but all possible excited states with the same quantum numbers. If we denote the normalized state taken from applying our nucleon operator on the vacuum as $|N\rangle$, then the Euclidean three-point correlator may be expressed as

$$\langle N(t_{\text{sep}})J_\mu(\tau)N(0) \rangle = \sum_{m,n} \langle N|m \rangle \langle m|J_\mu|n \rangle \langle n|N \rangle e^{-E_m(t_{\text{sep}}-\tau)-E_n\tau}, \quad (3.19)$$

where two complete sets of state have been inserted with a normalization defined as $\langle m|m \rangle = 1$ and the source has been shifted to the origin. Since excited state energies are higher than the ground state, the corresponding exponentials decay faster than the ground state one. Therefore, in theory one may look at time separations large enough to filter out unwanted contributions and isolate the desired matrix element. In practice however this is difficult due to increasing gauge noise. A famous expression for the approximate signal-to-noise ratio of a nucleon (or any baryon) is

$$\rho = e^{-(m_N - \frac{3}{2}m_\pi)t_{\text{sep}}}. \quad (3.20)$$

From (3.20), it is clear that the signal-to-noise ratio of the nucleon drops off very quickly. Therefore, the lattice data must be fitted to remove the excited state contamination at time separations short enough in which there is a reasonable signal. The following subsections describe various fitting methods to accomplish this. We alert the reader that our examination of fitting models consider only zero-momentum transfer matrix elements. All fitting forms introduced below other than the plateau method may be generalized to

non-zero momentum transfer, however we do not give the general formulas here for sake of simplicity and because only zero-momentum transfer matrix elements are analyzed in this work.

3.3.1 Plateau method

We begin the survey of fitting methods with the most basic one, fitting to a constant. This technique can only be applied to zero-momentum transfer matrix elements. This comprises of inspecting the data as a function of time separation and only including points in which the decaying tails from excited state have fallen off, resulting in a plateau. The remaining data are fit to a constant. This technique has several issues. As mentioned previously, for many Lattice calculations, the time separations for which a credible plateau could be theoretically reached are too large and gauge noise prevents a clean signal from being extracted in those regions. Furthermore, excited state contamination from hadron three-point correlators depends on both the source-sink time separation and the current insertion separation between source and sink. Thus it is possible to observe a plateau with respect to one of these times but not the other, causing a possible unaccounted systematic. Details of this dependence will be worked through in the next subsection.

One of the main advantages of the plateau method is its simplicity. Often, a constant fit is not even needed and the matrix element may be extracted visually with this technique. Hence the plateau method is frequently used as a quick first check to obtain rough features of lattice data, even if it is not one of the final methods quoted in formal results. Additionally, extremely noisy correlators may not encompass errors small enough to resolve excited state contamination well, even at short time separations. For such data, other fitting methods often fail to converge, leaving the plateau method as the only option for low statistics studies. Such can be the case for preliminary studies of disconnected

diagrams. We remark that results from this thesis are accurate enough for resolution of excited states and this method is not quoted in the Selected Results chapter.

3.3.2 Two-state model

The most straight forward way of accounting for excited state contamination and removing it is a fit based on the two-state model. Here, we assume that the main contribution from excited states comes from only the first excited state. We use the notation of $c^{2pt}(t_{\text{sep}})$ and $c^{3pt}(t_{\text{sep}}, \tau)$ for the two and three-point correlators respectively, where t_{sep} is the source-sink separation, defined as $t_{\text{sep}} = t - t_0$ and τ is the loop insertion time. Note that c^{3pt} and c^{2pt} only have two and one arguments respectively because they have been defined with t_0 , the source time, shifted to the origin. With these coordinates the three-point function can be viewed as a lower triangular matrix with the end points removed, or that $0 < \tau < t_{\text{sep}}$.

The two-point correlator is given by

$$c^{2pt}(t_{\text{sep}}) = Z_1 e^{-E_1 t_{\text{sep}}} + Z_2 e^{-E_2 t_{\text{sep}}} + \dots, \quad (3.21)$$

where Z_1, Z_2, \dots are the overlap factors of the interpolating field with the ground and excited states. In our notation, a subscript 1 refers to the ground state and 2+ corresponds to excited states. Using a similar notation, the three-point function is

$$\begin{aligned} c^{3pt}(t_{\text{sep}}, \tau) &= Z_1 \langle 1 | J_\mu | 1 \rangle e^{-E_1 t_{\text{sep}}} + \sqrt{Z_1 Z_2} \langle 1 | J_\mu | 2 \rangle e^{-E_1(t_{\text{sep}} - \tau)} e^{-E_2(\tau - t_0)} \\ &+ \sqrt{Z_2 Z_1} \langle 2 | J_\mu | 1 \rangle e^{-E_2(t_{\text{sep}} - \tau)} e^{-E_1(\tau - t_0)} + Z_2 \langle 2 | J_\mu | 2 \rangle e^{-E_2 t_{\text{sep}}} + \dots \end{aligned} \quad (3.22)$$

Dividing the three-point correlation function by the two-point correlator and only

looking at the real part of the ratio yields a compact fitting form with 4 parameters.

$$R(t_{\text{sep}}, \tau) = \frac{c^{3pt}(t_{\text{sep}}, \tau)}{c^{2pt}(t_{\text{sep}})} \approx A + Be^{-\Delta E\tau} + Be^{-\Delta E(t_{\text{sep}}-\tau)} + Ce^{-\Delta Et_{\text{sep}}}. \quad (3.23)$$

This functional form captures the t_{sep} and τ dependence of the data. This same strategy may be applied to higher excited states to form a three-state or even four-state model. However, this increases the number of fit parameters and the fits typically become unstable with more than three exponentials due to gauge noise. Even the two-state model may sometimes prove ineffective if the data are too noisy.

3.3.3 Exponential Fit

Another fitting method is to combine the plateau technique with the two-state model. One of the weaknesses of the plateau method is due to introducing a potential systematic from the data plateauing in the τ dependence, but not in t_{sep} . A way around this is to do a fit similar to the two-state model, however only taking the middle current insertion time.

From (3.23), it is clear that at $\tau = \frac{t_{\text{sep}}}{2}$, excited state contamination from the $\langle 1|J_\mu|2 \rangle$ and $\langle 2|J_\mu|1 \rangle$ contributions is at a minimum. Therefore, instead of utilizing three exponentials, one may assume that at middle insertion times the data have plateaued in τ dependence and only a single exponential fit is needed to capture the remaining t_{sep} dependence.

$$R(t_{\text{sep}}) \approx A + Ce^{-\Delta Et_{\text{sep}}}, \quad (3.24)$$

Only $\tau = \frac{t_{\text{sep}}}{2}$ points are inserted into (3.24). While this does not use all the data

points, as in (3.23), the fitting form is more stable at small differences of t_{sep} because fewer fitting parameters are present and only a single exponential is fitted. In subsequent chapters we refer to this method as the “exponential fit”.

3.3.4 Summation method

The summation method is another alternative that addresses the same issues as the exponential fit. Just as in the previous subsection, this procedure only involves fitting to a single exponent. Unlike the “exponential fit” however, the summation technique still uses data points from all τ . If we define a modified ratio as

$$\tilde{R}(t_{\text{sep}}) := \sum_{\tau=1}^{t_{\text{sep}}-1} \frac{\langle c^{3pt}(t_{\text{sep}}, \tau) \rangle}{c^{2pt}(t_{\text{sep}})}, \quad (3.25)$$

then the correlators may be expanded under this ratio which approximately gives

$$\tilde{R}(t_{\text{sep}}) \approx A + t_{\text{sep}}(\langle 1|J\mu|1 \rangle + Be^{-\Delta Et_{\text{sep}}}). \quad (3.26)$$

Therefore, if $\tilde{R}(t_{\text{sep}})$ and $\tilde{R}(t_{\text{sep}} - 1)$ are subtracted, the difference is proportional to the matrix element of interest plus a single exponential.

$$\tilde{R}(t_{\text{sep}}) - \tilde{R}(t_{\text{sep}} - 1) \approx A + Be^{-\Delta Et_{\text{sep}}} \quad (3.27)$$

This provides the same benefits as the exponential fit although it uses all of the data. Systematics of the summation method may potentially vary more than with the exponential fit however, since multiple τ points are included without explicitly tracking their dependence. Nevertheless, the summation method is another fantastic alternative to the two-state model. This concludes our discussion of different fitting forms to reduce excited state contamination. These techniques will be employed to extract the ground

state of disconnected matrix elements in Chapter 6.

CHAPTER 4

Introduction to Numerical Linear Algebra for Sparse Matrices

This chapter familiarizes the reader with the basics of numerical linear algebra, with particular focus on linear solvers and trace estimators for sparse matrices. No knowledge of quantum field theory or even advanced physics is required for this and the subsequent chapter. This examination of numerics is completely self-contained, although references and parallels to our physics application will be made throughout.

The first section of this chapter presents direct methods, iterative methods, and the concept of numerical stability. Subsequently, we study iterative linear solvers and Krylov subspace methods, such as conjugate gradient. Additionally, the notion of a preconditioner is established with adaptive algebraic multigrid as an example. Lastly, we review the basic stochastic trace estimator and possible improvements, ultimately leading to hierarchical probing.

4.1 Preamble

We start by defining quintessential concepts in numerical linear algebra. These include direct vs iterative methods, the conditioning of a problem, and numerical stability. This section contains many useful keywords, however it also builds the basic logic of numerical techniques which will be expanded on in the forthcoming sections

4.1.1 Direct vs iterative methods

We broadly define two classes of numerical algorithms, direct and iterative. Direct methods calculate the solution of a given problem in a finite number of steps. Such methods compute an exact solution, modulo floating point error. Iterative techniques, on the other hand, are in general not expected to obtain an exact answer in a finite number of steps.

Iterative methods typically take an initial guess as input and make a progressive set of approximate solutions, where each solution depends only on the previous one. In general, even with infinite precision arithmetic, an iterative solution becomes exact only in the limit of infinite steps. Therefore a convergence criterion is used to terminate the algorithm after sufficient accuracy is reached. A prototypical measure of convergence is the absolute difference, or residual.

A simple example demonstrating the difference between direct and iterative algorithms is shown below. Consider the equation

$$2x^3 - 24 = 30. \tag{4.1}$$

A direct technique to solve for x is as follows.

We apply Algorithm 1 to (4.1), of course these are the same algebraic steps one would use to solve such an equation by hand.

Algorithm 1 Directly solve $ax^3 + b = c$

1: $d = c - b$

2: $d = \frac{d}{a}$

3: $x = \sqrt[3]{d}$

1. $2x^3 = 54$

2. $x^3 = 27$

3. $x = 3$

In the last step we have only given purely real solutions for x . Notice that being able to take the cube-root of a number is required for this method, if this operation is unavailable this technique cannot be used. An alternative that does not require the cube-root is the following iterative algorithm.

Algorithm 2 Iteratively solve $ax^3 + b = c$

1: $d = c - b$

2: $i = 0$

3: $x_0 = \text{initial guess}$

4: **while** $ax_i^3 - d > tol$ **do**

5: $i = i + 1$

6: $x_i = x_{i-1} - \frac{1}{3} \left(x_{i-1} - \frac{d}{ax_{i-1}^2} \right)$

7: **end while**

As discussed previously, this method takes a starting value and tolerance as input. The starting value, x_0 is passed as an initial guess and the tolerance, tol is used as the stopping criterion. The iteration of x_i only depends on x_{i-1} , therefore only a single number needs to be stored. We set $x_0 = 1$ and show the first 7 iterations for (4.1).

0. $x_0 = 1.0000$
1. $x_1 = 9.6667$
2. $x_2 = 6.5408$
3. $x_3 = 4.5709$
4. $x_4 = 3.4780$
5. $x_5 = 3.0627$
6. $x_6 = 3.0013$
7. $x_7 = 3.0000$

Algorithm 2 is a specific case of the Newton-Raphson method, which provides an iterative approach to solving for the roots of any differentiable function.

$$x_i = x_{i-1} - \frac{f(x_{i-1})}{f'(x_{i-1})} \quad (4.2)$$

Although powerful, Newton's method has its own setbacks such as required knowledge of the derivative and divergence near a stationary point. The advantages and disadvantages of direct versus iterative methods in real-world applications are more complex than in our toy example. In Lattice QCD, the massive dimension of the Dirac matrix constrains computation of point-to-all propagators to be performed with iterative techniques. This will be elaborated on in the following section.

4.1.2 Condition number and numerical stability

Before exploring common iterative methods used in Lattice QCD we must understand the role of conditioning. The conditioning of a problem relates a function and variation in an input argument to the change of the output. This is useful in determining how the error in an input parameter can affect the error in the final result of a calculation. A problem is considered “well-conditioned” if its condition number is low, conversely the problem is “ill-conditioned” if its condition number is high.

In Lattice QCD, a common task is inverting the Dirac operator on a source vector: $Ax = b$. We caution the reader about switching to standard linear algebra notation, A is any matrix (in our case the Dirac matrix), b is the input vector (source), and x is the solution (quark propagator). The ensuing discussion is relevant for any matrix, although our specific interest is in the Lattice Dirac operator. If the condition number of a matrix A is large, any small inaccuracy in knowledge of b can cause large error in x . Conversely, if the condition number of A is small, the error on x will be comparable to the error of b .

We arrange the singular values of A from smallest to largest and use the notation $\sigma_1(A) \leq \sigma_2(A) \leq \dots \sigma_N(A)$. The condition number for a matrix A is consequentially defined as

$$\kappa_2(A) = \frac{\sigma_N(A)}{\sigma_1(A)}, \quad (4.3)$$

or the ratio of maximal and minimal singular values. The κ_2 indicates that we have defined the condition number through the 2-norm, while other definitions exist, (4.3) is one of the most widely used. Matrices with large κ_2 are considered ill-conditioned, while those with a low κ_2 are well-conditioned. Matrices that have zero as a singular value have infinite condition number and are of course not invertible.

Another influence on the error of x is numerical stability. While condition number is

the property of a matrix, or more generally of a particular problem, stability is a property of the algorithm employed to solve that problem. There may be many possible algorithms to solve a specific problem. The combination of the algorithm and problem determine the behavior of the error on the solution.

A useful way to analyze the stability of an algorithm is through a technique called backward error analysis. Let us revisit the problem of inverting a matrix on an input vector: $Ax = b$. We call the exact solution: x , and the approximate solution given by a particular algorithm: x' . The forward error is simply the absolute value of the difference between the exact and approximate solutions: $|x' - x|$. Conversely, backward error is the smallest Δb for $Ax' = b + \Delta b$. Put differently, backward error is a measure of how different of a problem the algorithm solved compared to the desired one, assuming infinite precision arithmetic.

An algorithm is considered to be backward stable if $\frac{|x' - x|}{|x|} \leq \kappa_2(A) \frac{|\Delta b|}{|b|}$. In contrast, an algorithm is thought of as unstable if $\frac{|x' - x|}{|x|} > \kappa_2(A) \frac{|\Delta b|}{|b|}$. Backward stability ensures that if the condition number of the problem is of order unity, small errors in the input will lead to small errors in the output. A well conditioned problem and stable algorithm suggest that inaccuracies in the solution are from errors in the input, not from propagation of floating point error.

4.2 Linear solvers

Now we turn our attention to algorithms for solving a linear system with the Dirac matrix for a particular vector $Ax = b$. In this section we will cover common iterative algorithms that are used to solve linear systems, including conjugate gradient and Bi-CGStab. Additionally, the notion of Krylov subspace methods and preconditioners will be introduced. We will conclude with the basic outline of adaptive algebraic multigrid, the

state of the art preconditioner to a linear solver, which is employed ubiquitously in Lattice QCD.

4.2.1 Conjugate Gradient

Just as in our toy model, both direct and iterative methods exist for inverting the Dirac matrix on a source. A common direct algorithm to solve a linear system of equations is Gauss-Jordan elimination. Through this technique, the solution is reached by manipulating the rows of A in a systematic way. The allowed operations are multiplying one row by any number and adding a multiple of one row to another.

Gauss-Jordan elimination is often extremely useful for solving small linear systems by hand. Additionally, the algorithm is straightforward to implement computationally and can handle the inversion of moderately sized matrices on a vector. For matrices of dimension tens or hundreds of millions, such as in Lattice QCD, Gauss-Jordan elimination becomes completely impracticable due to its $O(N^3)$ scaling. Just as in our root-finding toy example, an alternative is to use an iterative method, which will require an initial guess and tolerance. One example of the aforementioned algorithm is conjugate gradient [43].

Conjugate gradient is an algorithm for obtaining the solution of a linear system in the case of a Hermitian and positive-definite matrix. Although the Dirac operator is not Hermitian, this method may still be applied by working on the normal equation.

$$A^\dagger Ax = A^\dagger b \tag{4.4}$$

We begin the examination of the conjugate gradient algorithm with the notion of conjugate vectors. Two vectors u and v are considered to be conjugate with respect to A if

$$u^\dagger Av = 0. \quad (4.5)$$

Since A is Hermitian, if u is conjugate to v , v is also conjugate to u .

$$u^\dagger Av = v^\dagger Au = 0 \quad (4.6)$$

Consider N vectors p_1, p_2, \dots, p_N all conjugate to each other, where N is the dimension of A . We may express the solution of $Ax = b$ using these vectors as a basis, with corresponding coefficients α_i .

$$x = \sum_i^N \alpha_i p_i \quad (4.7)$$

Therefore,

$$Ax = \sum_i^N \alpha_i Ap_i \quad (4.8)$$

$$p_j^\dagger Ax = \sum_i^N \alpha_i p_j^\dagger Ap_i \quad (4.9)$$

$$p_j^\dagger Ax = \alpha_j p_j^\dagger Ap_j, \quad (4.10)$$

where in the last step the sum collapses to one term due to all p_i being mutually conjugate to each other. Furthermore, since $Ax = b$, only knowledge of p_j is required to solve for α_j .

$$p_j^\dagger b = \alpha_j p_j^\dagger Ap_j \quad (4.11)$$

$$\alpha_j = \frac{p_j^\dagger b}{p_j^\dagger Ap_j} \quad (4.12)$$

To summarize, in order to find the exact solution x , we must obtain N conjugate vectors p_j and compute the respective coefficients α_j . For matrices of large dimension however, this method is prohibitively expensive, similar to Gauss-Jordan elimination. In fact, so far the algorithm we have presented is a direct method, not an iterative one. It should be noted that some p_j will be more important than others in determining an approximate solution x' . Therefore, we may convert this algorithm into an iterative one by starting with one p_j and expanding the basis until x' is close to x within an input tolerance.

Akin to our toy example, this iterative algorithm will start with an initial guess x_0 . Next, consider the function

$$\phi(x) = \frac{1}{2}x^\dagger Ax - x^\dagger b. \quad (4.13)$$

The gradient of $\phi(x)$ is $Ax - b$, therefore the solution to the linear system of equations is simply the minimization of $\phi(x)$. This implies the first basis vector should be the negative gradient of $\phi(x)$ at x_0 .

$$p_0 = b - Ax_0 \quad (4.14)$$

Note that all following p_j will be conjugate to this gradient, thus the name conjugate gradient. We define the j th residual as

$$r_j = b - Ax_j. \quad (4.15)$$

The following p_j should be conjugate to all the previous ones and be built out of the current residual. This can be achieved with vector projections in an analogous manner to orthogonalization techniques.

$$p_{j+1} = r_j - \sum_{i=0}^j \beta_{i,j} p_i \quad (4.16)$$

$$\beta_{i,j} = \frac{p_i^\dagger A r_j}{p_i^\dagger A p_i} \quad (4.17)$$

Subsequently, the iteration is repeated with the successive approximate solution x_j

$$\alpha_j = \frac{p_j^\dagger b}{p_j^\dagger A p_j} = \frac{p_j^\dagger (r_{j-1} + A x_{j-1})}{p_j^\dagger A p_j} = \frac{p_j^\dagger r_{j-1}}{p_j^\dagger A p_j} \quad (4.18)$$

$$x_j = x_{j-1} + \alpha_j p_j, \quad (4.19)$$

where the final step of (4.18) is taken by recognizing that x_{j-1} is made up of a linear combination of p_i that are all conjugate to p_j . Furthermore, in Appendix C we show that all $\beta_{i,j} = 0$ for $i < j$, therefore only the last term of the sum survives in (4.16). We relabel $\beta_{j,j}$ as β_j and present the full algorithm below.

Algorithm 3 Conjugate Gradient

- 1: $x_0 =$ initial guess
 - 2: $j = 0$
 - 3: $r_0 = b - A x_0$
 - 4: $p_1 = r_0$
 - 5: **while** $\frac{\|r_j\|}{\|b\|} > tol$ **do**
 - 6: $j = j + 1$
 - 7: $\alpha_j = \frac{r_{j-1}^\dagger r_{j-1}}{p_j^\dagger A p_j}$
 - 8: $x_j = x_{j-1} + \alpha_j p_j$
 - 9: $r_j = r_{j-1} - \alpha_j A p_j$
 - 10: $\beta_j = \frac{r_j^\dagger r_j}{r_{j-1}^\dagger r_{j-1}}$
 - 11: $p_{j+1} = r_j + \beta_j p_j$
 - 12: **end while**
-

Instead of applying an additional matrix-vector multiplication, as in (4.15), the current

residual in Algorithm 3 is maintained iteratively. Additionally, the stopping criterion is the ratio of $\frac{\|r_j\|}{\|b\|}$, or the relative residual. Details of the re-expression of α and β can be found in Appendix C.

More generally, conjugate gradient is in a class of algorithms denoted as Krylov subspace methods. An order r Krylov subspace is a vector space spanned by the vectors $b, Ab, A^2b, \dots, A^{r-1}b$. The conjugate gradient method may be viewed as an algorithm that quickly constructs the optimal polynomial to approximate x using this space.

4.2.2 Bi-conjugate gradient stabilized method

An extension of conjugate gradient that allows A to be non-symmetric is the bi-conjugate gradient method. Under this algorithm, two sets of search directions and residual vectors are iterated over. Further modifications yield improvement in convergence behavior, the so-called Bi-CGStab method [43, 44]. The recurrence relations in Bi-CG are

$$p_j = r_{j-1} + \beta_j p_{j-1} \tag{4.20}$$

$$\hat{p}_j = \hat{r}_{j-1} + \beta_j \hat{p}_{j-1} \tag{4.21}$$

$$r_j = r_{j-1} - \alpha_j A p_j \tag{4.22}$$

$$\hat{r}_j = \hat{r}_{j-1} - \alpha_j A^\dagger \hat{p}_j, \tag{4.23}$$

where α_j and β_j are defined as

$$\alpha_j = \frac{\rho_j}{\hat{p}_j^\dagger A p_j} \quad (4.24)$$

$$\beta_j = \frac{\rho_j}{\rho_{j-1}} \quad (4.25)$$

$$\rho_j = \hat{r}_{j-1}^\dagger r_{j-1}. \quad (4.26)$$

These recurrences can be done implicitly however, we encourage the reader to look at [43] for details and simply provide pseudocode of the algorithm.

Algorithm 4 Bi-CGStab

```

1:  $x_0 =$  initial guess
2:  $j = 0$ 
3:  $r_0 = b - Ax_0$ 
4:  $\hat{r}_0 = r_0$  (not a unique choice)
5:  $\rho_0 = \alpha_0 = \omega_0 = 1$ 
6:  $v_0 = p_0 = 0$ 
7:  $s = r_0$ 
8: while  $\frac{\|s\|}{\|b\|} > tol$  do
9:    $j = j + 1$ 
10:   $\rho_j = \hat{r}_0^\dagger r_{j-1}$ 
11:   $\beta_j = \frac{\alpha_j \rho_j}{\omega_{j-1} \rho_{j-1}}$ 
12:   $p_j = r_j + \beta_j(p_j - \omega_{j-1} v_{j-1})$ 
13:   $v_j = Ap_j$ 
14:   $\alpha_j = \frac{\rho_j}{\hat{r}_0^\dagger v_j}$ 
15:   $s = r_{j-1} - \alpha_j v_j$ 
16:   $t = As$ 
17:   $\omega_j = \frac{t^\dagger s}{t^\dagger t}$ 
18:   $x_* = x_{j-1} + \alpha_j p_j$ 
19:   $x_j = x_* + \omega_j s$ 
20:   $r_j = s - \omega_j t$ 
21: end while

```

In Algorithm 4, if the relative norm of s is small enough, x_* is returned as the solution. Bi-CGStab requires two-matrix vector (matvec) applications and four inner products per iteration, unlike conjugate gradient which merely needs one matvec and one inner product

per iteration. Conjugate gradient and Bi-CGStab are simply two examples of Krylov subspace based linear solvers, many more exist, including MR, GMRES, and GCR to name a few [45].

4.2.3 Preconditioners and multigrid

An approach to speedup the convergence of an iterative solver is to solve a preconditioned system, or

$$M^{-1}Ax = M^{-1}b, \tag{4.27}$$

where M is known as a preconditioner. If M is easy to invert and $M^{-1}A$ has a lower conditioner number than A , the solution x may be computed with less computational cost than the unpreconditioned system. In general, there must be a balance between the effectiveness and ease of cost of the preconditioner. The most inexpensive preconditioner is the identity $M = I$, since $M^{-1} = I$, however this is identical to the original, unpreconditioned system. Alternatively, the most effective preconditioner is $M = A$, since this would give a condition number of 1 for $M^{-1}A$, however this requires computing A^{-1} , which is even more difficult than the original problem of inverting A on a single input vector.

A preconditioner that has received tremendous success in Lattice QCD is adaptive algebraic multigrid [46], motivated by the Dirac operator's property of local coherence. This is the property that the low lying eigenmodes of the Dirac matrix tend to be approximately localized. We do not give a full rigorous treatment of adaptive multigrid here and simply outline the motivation and basic steps. The key idea is to reduce error from slow to converge, low frequency modes by coarsening A and inverting with a smaller, coarse operator A_c . The low frequencies from the fine matrix will act like higher frequency modes on A_c . After inverting A_c , this coarse grid correction is transferred and added back to the

current fine solution. The beauty of multigrid is that this process is recursive, allowing many levels of coarsening until the cost of inverting the coarsest matrix is near negligible in principle.

The two main features of any multigrid algorithm are the coarse grid linear solver, or “smoother”, and transfer operators that prolongate and restrict between fine and coarse levels. The transfer operators may be thought of as projectors that either convert a fine vector to a coarse one, or vice versa. In Lattice QCD, the near-null space components of the Dirac operator slow down the convergence of Krylov solvers such as conjugate gradient or Bi-CGStab. Therefore, the coarse Dirac matrix is constructed out of adaptively found near-null space vectors, which are rich in low modes. These near-null vectors are chopped up or “aggregated” into blocks. Since the Dirac matrix has the property of local coherence, these blocks are able to represent the contribution from hundreds or even thousands of low-lying eigenvectors.

We outline the setup phase of a two-level adaptive multigrid algorithm below, going from determining the near-null vectors to constructing the coarse Dirac operator.

Algorithm 5 Adaptive MG Setup

- 1: n_v = number of null vectors
 - 2: **for** $j = 1 : n_v$ **do**
 - 3: Solve $Ax = 0$ for a random starting guess vector with either a small, fixed number of iterations or relaxed tolerance.
 - 4: **end for**
 - 5: Aggregate v_j into blocks and orthonormalize them (typically done with a technique called QR factorization).
 - 6: Construct prolongator P (columns of blocked and orthonormal v_j).
 - 7: Build coarse operator $A_c = P^\dagger AP$.
-

Here we have used that the restricting transfer operator $R = P^\dagger$. This is a convenient choice that maintains γ_5 hermiticity on coarse grids. To expand multigrid to more than two levels, the setup in Algorithm 5 is repeated for each coarse level. Following this setup,

the solver typically follows a so-called V-cycle of inverting on the fine grid, then inverting on coarser grids, and finally working back up and inverting on the finest grid again. It is worth mentioning that the multigrid setup phase may be costly, however with enough source vectors as discussed in Chapter 3, this cost is easily amortized. For computing disconnected observables, the number of vectors that need to be solved far outweighs the setup cost. For relevant lattice parameters, multigrid provides an order of magnitude speedup compared to other methods [46, 47, 48].

4.3 Trace estimation techniques

At the end of Section 3.2, the difficulty of disconnected diagrams and need to stochastically compute the all-to-all propagator was covered. Here, we explore the details of relevant algorithms to estimate a quark loop, put simply, the trace of the inverse of a large, sparse matrix. Although a few optimizations that are most useful in the case of the matrix inverse are introduced, the basic techniques that are presented in this section are applicable to the trace of any matrix function.

This discussion is premiered with the most widely used and standard method: Hutchinson’s trace estimator. This involves Monte Carlo averaging over matrix quadratures with random noise vectors. Subsequently, we describe probing, a deterministic technique which is designed to suppress error from nearest neighbors. This is followed with a brief introduction to noise dilution, a method that splits the vector space of the trace estimator into smaller sub-spaces. Finally, we present an advanced trace estimator that was developed at William and Mary: hierarchical probing. To distinguish between different types of probing in the following text, we refer to ordinary probing as “classical probing”.

4.3.1 Hutchinson trace

The Hutchinson method capitalizes on the fact that $\text{Tr}[f(A)] = E[z^\dagger f(A)z]$, put differently, the trace is equal to the expectation value of quadratures of $f(A)$ with random noise vectors. The term $f(A)$ can be any function of A , our application focuses on $f(A) = A^{-1}$. Therefore, the following Monte Carlo averaging with s random vectors yields an unbiased estimator of the trace inverse.

$$\text{Tr}(A^{-1}) = \frac{1}{s} \sum_{j=1}^s z_j^\dagger A^{-1} z_j \quad (4.28)$$

The vectors z_j are made up of N random variables that satisfy $E(z_j(k)z_j(k')) = \delta_{kk'}$. It is important to note that for each step of this estimator, $A^{-1}z_j$ must be computed, therefore all technology from the previous section is commissioned in each step of this procedure. The error of this estimator is the usual $\sqrt{\text{Var}[\text{Tr}(A^{-1})]}/s$ of a Monte Carlo process. If the variance is large, the estimator suffers from poor convergence. Substantial investigation has gone into finding random vectors that minimize this variance [49, 50, 51]. For Gaussian distributed noise, this variance is proportional to the sum of the absolute square of all the matrix elements of A^{-1} .

$$\text{Var}[\text{Tr}_{\text{Gaussian}}(A^{-1})] = 2\|A^{-1}\|_F^2 = 2 \sum_{i,j}^N |A_{i,j}^{-1}|^2. \quad (4.29)$$

We define Rademacher vectors or \mathbb{Z}_2 noise, as vectors with ± 1 uniformly distributed elements. Similarly, \mathbb{Z}_4 vectors have components uniformly distributed from $\pm 1, \pm i$. The respective variances are given by

$$\begin{aligned}
\text{Var}[\text{Tr}_{\mathbb{Z}_2}(A^{-1})] &= \|A^{-1}\|_F^2 - \sum_{i=1}^N |A_{i,i}^{-1}|^2 + \sum_{i \neq j}^N A_{i,j}^{-1} A_{j,i}^{-1*} \\
&= \sum_{i \neq j}^N (|A_{i,j}^{-1}|^2 + A_{i,j}^{-1} A_{j,i}^{-1*}) \tag{4.30}
\end{aligned}$$

$$\text{Var}[\text{Tr}_{\mathbb{Z}_4}(A^{-1})] = \|A^{-1}\|_F^2 - \sum_{i=1}^N |A_{i,i}^{-1}|^2 = \sum_{i \neq j}^N |A_{i,j}^{-1}|^2. \tag{4.31}$$

Notice that in the case of \mathbb{Z}_2 noise, there is an extra term $\sum \sum_{i \neq j}^N A_{i,j}^{-1} A_{j,i}^{-1*}$ which is not present for the variance of \mathbb{Z}_4 noise. This term could be positive or negative in general, signifying that \mathbb{Z}_2 vectors may be more or less efficient than \mathbb{Z}_4 vectors, depending on the specific properties of A^{-1} . If the complex phases of $A_{i,j}^{-1}$ and $A_{j,i}^{-1}$ are uncorrelated, this extra term averages to zero when the sum is taken, giving approximately equal variance between \mathbb{Z}_2 and \mathbb{Z}_4 . Such is the case for the quark propagator in Lattice QCD.

As a brief aside, if A is real and symmetric, this extra sum doubles with the first term in (4.30) to give a variance equal to exactly twice of (4.31). In other words, for the real and symmetric case, the variance of Rademacher vectors is twice the sum of the square of the off diagonal matrix elements of A^{-1} . While this may seem uneconomical in comparison to \mathbb{Z}_4 noise at first glance, utilizing \mathbb{Z}_2 vectors avoids complex arithmetic all together, resulting in half the number of floating point operations and yielding the same net efficiency.

Focusing on the \mathbb{Z}_4 case for simplicity, we find the variance is smaller than Gaussian distributed noise since it only involves the off diagonal elements of A^{-1} . Although this basic method is powerful, based on (4.31), large off diagonal elements in A^{-1} stagnate the convergence of the estimator. In subsequent subsections we will reconnoiter algorithms which attempt to suppress the error from the largest off diagonal elements.

4.3.2 Probing

From (4.31), it is clear that the largest off diagonal elements contribute the most in slowing down convergence of the Monte Carlo estimator. For a large class of matrices, the elements of A^{-1} decay from the non-zero structure of A . In Lattice QCD, the elements of the Dirac inverse decay as a function of geometric distance on the lattice. Although the basic Hutchinson method is powerful, it is unable to take advantage of this decay. An algorithm that benefits from these decay properties and discovers the structure of A^{-1} is probing, or what we refer to as “classical probing” [52].

Classical probing (CP) works by deploying a specific set of mutually orthogonal vectors into the trace estimator in (4.28), without the $\frac{1}{s}$ normalization. These vectors are designed to line up with the non-zero elements of the matrix of interest in a way that delivers the trace. CP works well for sparse matrices, because the number of probing vectors needed to reveal the exact diagonal is small. For example, if A^{-1} is a tridiagonal matrix, only two vectors (shown below) are required to remove all error and return the trace exactly.

$$\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ \vdots \end{pmatrix}$$

For a dense matrix, there is no sparsity pattern to profit from and CP is not effective. In Lattice QCD, A^{-1} is dense, but the aforementioned decay gives a structure to the relative magnitude of matrix elements which CP may exploit. This is achieved through the power series expansion of A^{-1} allowing the graph coloring of A^k to reveal the largest magnitude matrix elements of A^{-1} . The graph of A can be constructed with the nodes representing

the rows/columns of A and the matrix elements $A_{i,j}$ being the arcs that connects from nodes i to j . We assign a different color to each vertex that is within k hops of the nodes i and j and generate CP vectors that follow this coloring pattern. These vectors remove the largest error components in (4.31).

A judicious choice for the value of k must be made in order for CP to be beneficial. If the decay in A^{-1} is weak, many probing vectors are necessary to capture and remove significant error from the estimator. On the other hand, if the decay is fast, only a few CP vectors estimate the trace well and using more of them wastes computational resources. Furthermore, quadratures from a specific k coloring cannot be reused for higher distance coloring.

In Lattice QCD, one solution to the above issue is to generate a small number of probing vectors based on known structure. This could be probing with red/black partitions, spin/color components, timeslices, or any other known structure. Once this is done, s random noise vectors are created that follow the structure of the probing vectors. This is referred to as dilution [53, 54].

4.3.3 Hierarchical probing

Hierarchical probing (HP) is a method that has evolved from CP and dilution. HP approximates CP by discovering the structure of A^{-1} through geometry in a hierarchical way [55]. This is achieved by a nested coloring algorithm with the property that previous quadratures are reusable for higher distance coloring. The nested coloring scheme fully captures 2^{k+1} distances, however the coloring is suboptimal, causing more probing vectors to be used for a particular distance than in CP. These extraneous colors seep over and further annihilate error components of the trace estimator from larger distances.

For lattices of purely even powers of two, the HP basis is described by permuted

columns of the Hadamard matrix. For other dimensions, Fourier vectors may be used. We restrict the rest of the discussion of HP to Hadamard vectors.

Hadamard vectors are columns of an N dimensional square matrix with entries either $+1$ or -1 . All columns vectors of the Hadamard matrix are mutually orthogonal, a prerequisite for any probing basis. Furthermore, a recursion relation exists relating Hadamard matrices of different dimensions. If we label the Hadamard matrix of dimension N as H , the Hadamard matrix of order $2N$ is

$$\begin{bmatrix} H & H \\ H & -H \end{bmatrix}.$$

To demonstrate the power of the hierarchical nature of HP, let us revisit the example of a tridiagonally dominant matrix. The two vectors from the previous subsection eliminate error from alternating diagonal bands, including the two bands closest to the principal diagonal, thus greatly reducing the error of a tridiagonally dominant matrix. If the user decides to probe to a higher distance post factum, the new probing basis does not share any columns with the previous one.

$$\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ \vdots \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ \vdots \end{pmatrix}$$

Alternatively, the same cancellation of the diagonal bands can be attained by feeding two reordered Hadamard vectors into (4.28), with the $\frac{1}{s}$ normalization intact. Furthermore, to increase the probing distance, the original basis is extended by two additional vectors and previous work is not wasted.

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ \vdots \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \\ \vdots \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ \vdots \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \\ 1 \\ -1 \\ \vdots \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ \vdots \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \\ \vdots \end{pmatrix}$$

We remark that our example is equivalent to one dimensional probing and a few subtleties regarding the intelligence of the hierarchical coloring algorithm are not demonstrated from it. Moreover, the number of colors between CP and HP match exactly for a lattice of one dimension. We encourage the reader to follow [55], in which the recursive coloring algorithm is shown for toroidal lattices of arbitrary dimension. Remarkably, the appropriate permutations of the Hadamard matrix may be computed with simple bit-arithmetic on local sites.

Just like CP, nested coloring is a deterministic algorithm. To avoid this, we perform an elementwise product of every HP vector with a random \mathbb{Z}_4 vector, $z = z_0 \odot z_h$. This procedure preserves the relative sign between all components of the coloring basis and therefore does not disturb the desired cancellation properties at 2^{k+1} distances. The random noise does mix up residual contributions at higher distances that are not killed

from probing, thus yielding an unbiased estimator and providing a small boost of variance reduction to the original probing basis. This gives the interpretation of HP as a procedure that generates a nested probing basis from any starting vector.

For strange quark masses, HP has provided an order of magnitude improvement compared to the basic Hutchinson technique [55]. Furthermore, experiments up to 512 HP vectors with these matrices have illustrated an error that falls of as $\frac{1}{s}$, instead of the notorious $\frac{1}{\sqrt{s}}$ of the basic Monte Carlo method. This algorithm has been applied successfully in conjunction with spin/color and full time dilution to execute a high precision calculation of strange vector current and axial vector current form factors [26]. Nevertheless, it will be shown in the following chapter that HP provides little benefit for computations at light quark masses. This is due to the fact that A^{-1} becomes dominated by near null components that do not decay at light masses. A solution to this issue is to solve for the low lying subspace of A exactly and project these low modes out of A^{-1} . This is known as deflation.

CHAPTER 5

SVD Deflation as a Method of Variance Reduction For Estimating the Trace of a Matrix Function

In the previous chapter, we introduced the standard Monte Carlo trace estimation method, as well as improvements to the basic algorithm, such as dilution and hierarchical probing (HP). In this chapter, we cover in detail another improvement to the Hutchinson trace estimator and present novel contributions to both the fields of Numerical Analysis and Lattice QCD in understanding the efficacy of this method. This technique involves solving for part of the matrix exactly and projecting it out of the trace estimator, a scheme known as deflation. We specifically study a new variant of deflation based on singular value decomposition (SVD).

This chapter begins with an overview of SVD deflation and the difficulty in studying its effect on the variance of random noise trace estimation. Next, we introduce a model based on random matrix theory for qualitatively describing the effects of deflation on

variance. Only the spectrum of singular values is needed as input for the model. Specifics of this procedure are not given here, but can be found in [56]. Remarkably, the model predicts higher variance reduction for non-Hermitian matrices compared to Hermitian ones, a unique trait. We use the model to study various singular spectra in both the Hermitian and non-Hermitian case. Additionally, we test the effectiveness of the model's predictions on real world problems from a plethora of fields. A comparison of variance reduction between \mathbb{Z}_2 or \mathbb{Z}_4 vectors and Gaussian noise is also made. Finally, we outline the strategy of deploying our SVD deflation algorithm in Lattice QCD and demonstrate its synergy with HP on large gauge field configurations produced from production runs by the Jefferson Lab Lattice group.

5.1 SVD deflation and its impact on variance

In the previous chapter, the basic trace estimation algorithm by Hutchinson was introduced. A modification to this algorithm is to compute a partial singular value subspace of the matrix and remove it from the range of the operator. If we are interested in estimating $\text{Tr}(A^{-1})$, the lowest singular values of A should be computed and projected out. This yields an altered matrix, which behaves differently under the trace estimator. The goal is for the estimator to converge to the trace of this new operator more quickly than A^{-1} . Although a different trace than the desired one is computed, we can correct back to the original trace by adding the low subspace of A that was taken out of the stochastic estimator. This modifies the form of (4.28).

$$\text{Tr}(A^{-1}) = \frac{1}{s} \sum_{j=1}^s \left(z_j^\dagger A^{-1} z_j - z_j^\dagger V_1 \Sigma_1^{-1} U_1^\dagger z_j \right) + \text{Tr}(V_1 \Sigma_1^{-1} U_1^\dagger) \quad (5.1)$$

Above, the subscript 1 is used to describe the low singular triplets that have been

computed exactly and projected out. Let us now study the effects this removal has on the stochastic estimator. For simplicity, we consider the problem of calculating $\text{Tr}(A)$. Here, the largest singular triplets are the most important to remove. Our analysis will be general however; we are always free to set $A = D^{-1}$ post factum, where D is the Dirac matrix from Lattice QCD. Obviously, the smallest singular values of D are the largest of D^{-1} , furthermore A may be set to any function of D and the following discussion will hold without loss of generality.

Let U, V, Σ be the full singular triplets of A and U_1, V_1, Σ_1 be the k largest singular ones. If $U = (U_1, U_2), V = (V_1, V_2)$, and $\Sigma = \text{diag}(\Sigma_1, \Sigma_2)$, A can be decomposed as

$$A = U_1 \Sigma_1 V_1^\dagger + U_2 \Sigma_2 V_2^\dagger \equiv A_D + A_R, \quad (5.2)$$

and $\text{Tr}(A) = \text{Tr}(A_D) + \text{Tr}(A_R)$. $\text{Tr}(A_D)$ is the part that has been removed and $\text{Tr}(A_R)$ is the “remaining matrix” that is passed to the Hutchinson estimator. Because A_D is in some sense, the best k -rank approximation of A , one would expect that the variance on A_R will always be smaller than that of A . While this statement is correct for Gaussian distributed noise, non-intuitively, it is not always true for \mathbb{Z}_4 vectors. Let us analyze in depth (4.31), noting that we have cast our problem as estimating $\text{Tr}(A)$ not $\text{Tr}(A^{-1})$. Define the vectors $D = \text{diag}(A)$, $D_D = \text{diag}(A_D)$, and $D_R = \text{diag}(A_R)$ and the traceless matrices $\tilde{A} = A - \text{diag}(D)$, $\tilde{A}_D = A_D - \text{diag}(D_D)$, and $\tilde{A}_R = A_R - \text{diag}(D_R)$. It is well known from the properties of SVD that

$$\|A\|_F^2 = \sum_{i=1}^N \sigma_i^2 = \|\tilde{A}\|_F^2 + \|D\|_F^2, \quad \|A_R\|_F^2 = \sum_{i=k+1}^N \sigma_i^2 = \|\tilde{A}_R\|_F^2 + \|D_R\|_F^2. \quad (5.3)$$

Next we express the diagonals explicitly in terms of the elements of the unitary ma-

trices in SVD.

$$D(i) = \sum_{m=1}^N \sigma_m u_{im} \bar{v}_{im}, \quad D_R(i) = \sum_{m=k+1}^N \sigma_m u_{im} \bar{v}_{im}. \quad (5.4)$$

Above, \bar{v}_{im} is the conjugate of v_{im} . We also define $\Delta = (U \odot \bar{V})^\dagger (U \odot \bar{V})$, where \odot is the elementwise product of matrices, and \bar{V} is the elementwise conjugate of V . This gives,

$$\Delta_{ml} = \sum_{i=1}^N \bar{u}_{im} v_{im} u_{il} \bar{v}_{il}, \quad m, l = 1, \dots, N. \quad (5.5)$$

Following through, the expressions for the Frobenius norm of the diagonals in terms of Δ_{ml} are given below.

$$\begin{aligned} \|D\|_F^2 &= \sum_{i=1}^N (\sum_{m=1}^N \sigma_m \bar{u}_{im} v_{im}) (\sum_{l=1}^N \sigma_l u_{il} \bar{v}_{il}) = \sum_{m=1}^N \sum_{l=1}^N \sigma_m \sigma_l \sum_{i=1}^N \bar{u}_{im} v_{im} u_{il} \bar{v}_{il} \\ &= \sum_{m=1}^N \sum_{l=1}^N \sigma_m \sigma_l \Delta_{ml}. \end{aligned} \quad (5.6)$$

$$\|D_R\|_F^2 = \sum_{i=1}^N (\sum_{m=k+1}^N \sigma_m \bar{u}_{im} v_{im}) (\sum_{l=k+1}^N \sigma_l u_{il} \bar{v}_{il}) = \sum_{m=k+1}^N \sum_{l=k+1}^N \sigma_m \sigma_l \Delta_{ml}. \quad (5.7)$$

Therefore, the formula for the undeflated and deflated variances under \mathbb{Z}_4 noise may be represented as

$$\text{Var}[\text{Tr}_{\mathbb{Z}_4}(A)] = \sum_{m=1}^N \sigma_m^2 - \sum_{m=1}^N \sum_{l=1}^N \sigma_m \sigma_l \Delta_{ml} \quad (5.8)$$

$$\text{Var}[\text{Tr}_{\mathbb{Z}_4}(A_R)] = \sum_{m=k+1}^N \sigma_m^2 - \sum_{m=k+1}^N \sum_{l=k+1}^N \sigma_m \sigma_l \Delta_{ml}. \quad (5.9)$$

Due to the presence of the Δ_{ml} term from the diagonal, it is not guaranteed that

(5.9) will always be less than (5.8), in fact in the following sections we provide examples for which deflation hurts convergence of the trace estimator! It should be noted that for Gaussian vectors,

$$\text{Var}[\text{Tr}_{\text{Gaussian}}(A)] = \sum_{m=1}^N \sigma_m^2 \quad (5.10)$$

$$\text{Var}[\text{Tr}_{\text{Gaussian}}(A_R)] = \sum_{m=k+1}^N \sigma_m^2, \quad (5.11)$$

and the lack of the Δ_{ml} term ensures the variance of A_R will always be less than that of A . Furthermore, in the Gaussian case, the k best possible projection vectors to reduce variance are guaranteed to be the k largest singular triplets of A . The same is not necessarily true of \mathbb{Z}_2 or \mathbb{Z}_4 noise-based estimators.

In order to study the effects of SVD deflation with \mathbb{Z}_4 noise, random matrix theory with the Haar distribution was applied in [56] to give an expectation to Δ_{ml} . This produced concise formulas that only depend on the spectrum of the singular values of A . We split the results into two cases, one for non-Hermitian matrices and one for Hermitian ones. Both formulas are given below.

$$E(\text{Var}[\text{Tr}_{\mathbb{Z}_4, \text{non-Hermitian}}(A_R)]) = (N - k)\left(1 - \frac{1}{N}\right)(V_k + \mu_k^2) \quad (5.12)$$

$$E(\text{Var}[\text{Tr}_{\mathbb{Z}_4, \text{Hermitian}}(A_R)]) = (N - k)V_k\left(1 - \frac{1}{N + 1}\right) + k\frac{N - k}{N + 1}\mu_k^2 \quad (5.13)$$

Above, $\mu_k = \frac{1}{N-k} \sum_{m=k+1}^N \sigma_m$, and $V_k = \frac{1}{N-k} \sum_{m=k+1}^N (\sigma_m - \mu_k)^2$, respectively. Moreover, the ratio of the deflated to undeflated variances may be computed for the two cases. We do not derive the result here and simply take the solution from [56].

$$\frac{E(\text{Var}[\text{Tr}_{\mathbb{Z}_4, \text{non-Hermitian}}(A_R)])}{E(\text{Var}[\text{Tr}_{\mathbb{Z}_4, \text{non-Hermitian}}(A)])} = \frac{(N-k)(V_k + \mu_k^2)}{N(V_0 + \mu_0^2)} \quad (5.14)$$

$$\frac{E(\text{Var}[\text{Tr}_{\mathbb{Z}_4, \text{Hermitian}}(A_R)])}{E(\text{Var}[\text{Tr}_{\mathbb{Z}_4, \text{Hermitian}}(A)])} = \frac{\sum_{i=k+1}^N \sigma_i^2 - \frac{(N-k)^2}{N} \mu_k^2}{\sum_{i=1}^N \sigma_i^2 - N \mu_0^2} \quad (5.15)$$

In both cases, the faster the singular values grow, the higher variance reduction is expected to be achieved by deflation. It is clear that while the ratio in (5.14) is always less than 1, (5.15) can be greater than 1. Put differently, while deflation will never increase variance for non-Hermitian operators, it is possible for Hermitian ones. It is extraordinary to find a property for which non-Hermitian matrices perform better than Hermitian ones. The analysis in [56] however, shows that in general, non-Hermitian matrices are expected to have more variance to begin with.

5.2 The effect of the singular spectrum

Having factored out the contribution of the singular vectors with our model, we now study the effect of the singular value distribution using (5.12), (5.13), (5.14), and (5.15) to predict actual experiments. The larger the gap between deflated and undeflated singular values, the larger the expected reduction variance reduction. In the following experiments we study the effect of deflation for six different model distributions of σ_i .

Given a diagonal matrix of singular values Σ , we generate a pair of random unitary matrices U and V , and construct one Hermitian matrix $U\Sigma U^\dagger$ and one non-Hermitian matrix $U\Sigma V^\dagger$. For each model distribution we study matrices of several sizes. We report the ratio of the variance of the deflated matrix, where we deflate various percentages of its largest singular triplets, to the variance of the original undeflated matrix. This can be computed explicitly from $\|A\|_F^2$, or through our model. As statistically expected, beyond

small matrices of dimension less than 100, there is perfect agreement between our model's predictions and experimentally determined variances. Thus, we only present results from our model.

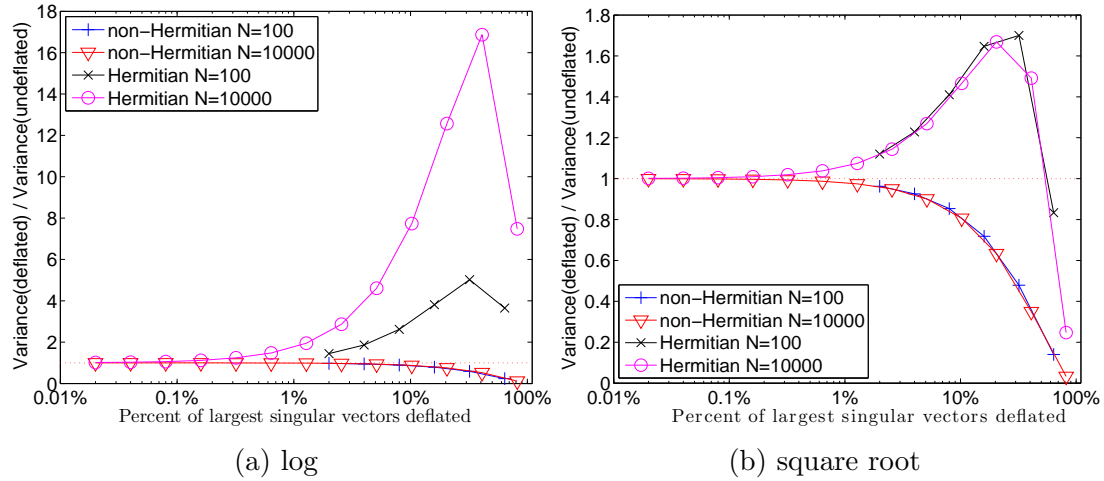


FIG. 5.1: On the left is a logarithmic spectrum: $\sigma_{N-i+1} = 1 + 2 \cdot \log(i)$. On the right is a square root spectrum: $\sigma_{N-i+1} = \sqrt{i}$. The dotted red line in both plots is a constant line at $y = 1$. Points below this line signify an improvement in variance with deflation. Points above the line denote a deflated operator with a higher Frobenious norm than the original matrix, a case in which deflation is hurtful and variance increases.

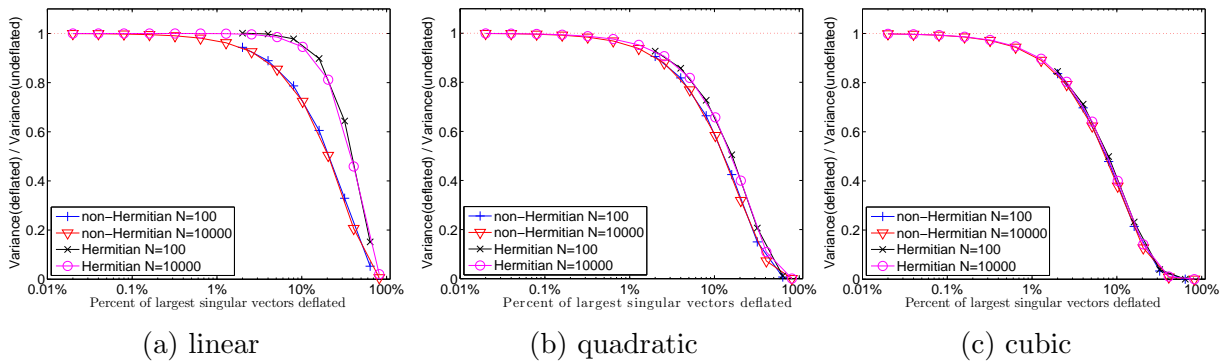


FIG. 5.2: Variance reduction ratios for matrices with spectra with linear, $\sigma_{N-i+1} = i$, quadratic, $\sigma_{N-i+1} = i^2$, and cubic, $\sigma_{N-i+1} = i^3$, growth rates.

In Figure 5.1a we consider a model where the singular values increase at a logarithmic rate with respect to their index. For Hermitian matrices the variance increases with the number of deflated singular triplets, and the problem is more pronounced with larger

matrix size. Although for non-Hermitian matrices the ratio is always below one, a substantial part of the spectrum must be deflated to reduce the variance appreciably. In Figure 5.1b the spectrum increases as the square root of the index, and the effects of deflation, although improved, still are not beneficial for Hermitian matrices.

In Figures 5.2a, 5.2b, and 5.2c the growth of the singular values is linear, quadratic, and cubic, respectively. The ratio is now below one for both types of matrices. We can see that with larger growth rates, the variance reduction is larger for a particular fraction of singular values deflated. Additionally, for sufficiently large growth rates, the difference between Hermitian and non-Hermitian matrices vanishes.

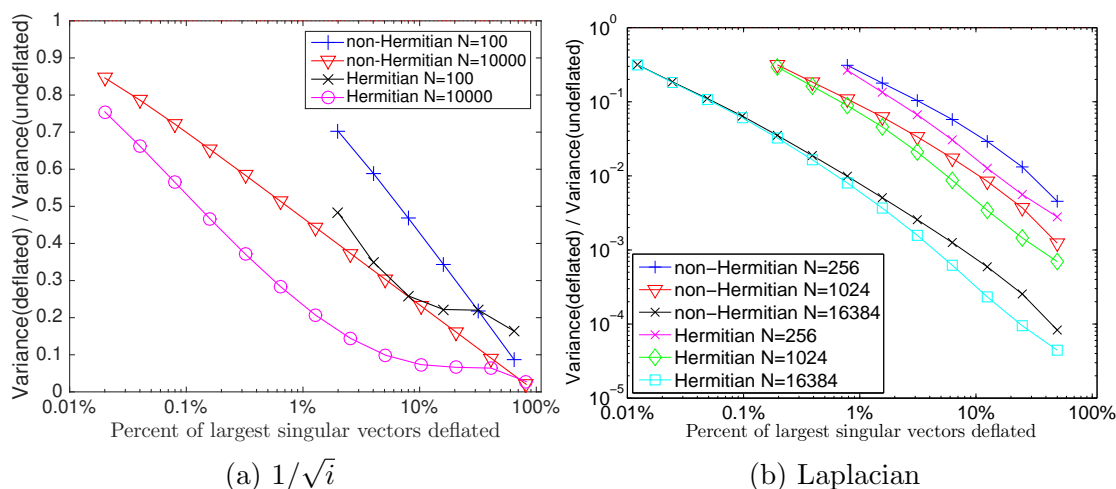


FIG. 5.3: On the left we have spectrum $\sigma_i = 1/\sqrt{i}$. The right plot shows the deflation of the inverse of a 2D discrete Laplacian on a grid $\sqrt{N} \times \sqrt{N}$ with Dirichlet boundary conditions.

For spectra that decay as a rational polynomial, the picture is different. Figure 5.3a shows an example where the spectrum is $\sigma_i = 1/\sqrt{i}$. There are a few large singular values that matter, but the rest do not reduce variance appreciably. The effect of this is that Hermitian matrices experience larger relative improvement with deflation over non-Hermitian matrices. We have observed this effect also for other rational polynomials, $1/i^p$, but the difference seems to peak at $1/\sqrt{i}$. This observation is particularly relevant to

Lattice QCD, in which the goal is to find the trace of the inverse of the Dirac matrix. In Figure 5.3b we study the spectrum of the inverse of the discrete Laplacian, a common problem that shares some features with the Dirac operator in the free field limit. We see significant variance reduction, especially as the lattice size grows. The Hermitian matrices continue to have an advantage over non-Hermitian matrices, but the practical difference is negligible.

5.3 Experiments on general matrices

The previous section studied the effect of spectra on matrices with random unitary singular vectors. In this section we investigate the extent to which our theory is applicable to general matrices with singular vectors that are not random. We choose four matrices from the University of Florida sparse matrix collection [57] with relatively small sizes (675–2000) that are derived from real world problems in various fields, such as chemical transport modeling and magnetohydrodynamics. In all the following results, deflation is studied on the estimator of $\text{Tr}(A^{-1})$. The discussion is split into two subsections, one continuing the analysis of \mathbb{Z}_4 noise as before and another studying the effects of SVD deflation on Gaussian distributed vectors.

5.3.1 Experiments with \mathbb{Z}_4 noise

Here, we continue the analysis of the previous section by applying our model of variance reduction under \mathbb{Z}_4 noise to four matrices from the University of Florida sparse matrix collection [57]. Three of the matrices are non-Hermitian and one is Hermitian, more strictly real and symmetric. We apply the appropriate formulas from the model for these cases.

For Figures 5.4a and 5.4b, the model and experimental results agree very closely. Both demonstrate dramatic variance reduction even when deflating a small fraction of the SVD

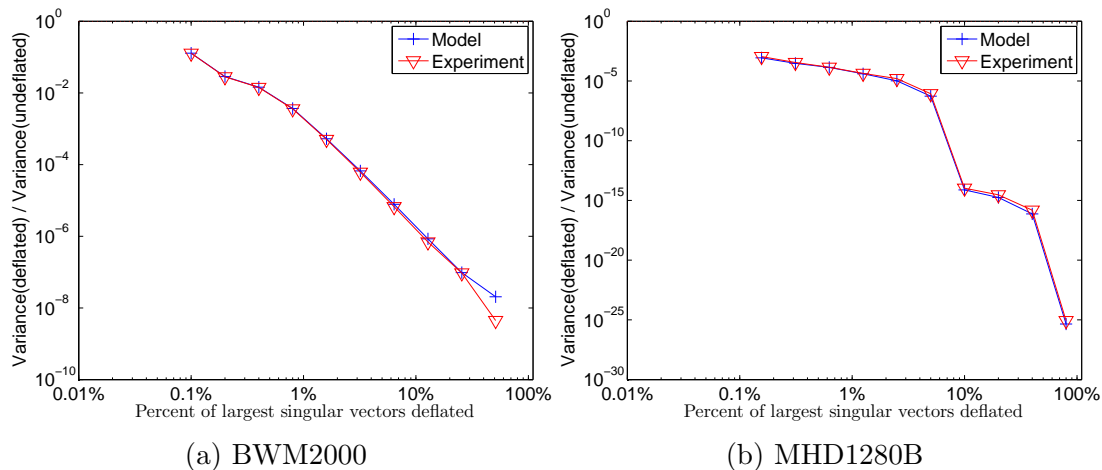


FIG. 5.4: Matrix BWM2000 has a size of $N = 2000$ and condition number of $2.37869e+5$. Matrix MHD1280B has $N = 1280$ and a condition number of $4.74959e+12$. Both matrices are real, non-symmetric.

space. Both matrices have a high condition number, implying that their smallest singular values contribute most of the variance to the inverse trace estimator. Hence it pays to remove them.

In contrast, deflation does not improve the variance for the matrix in Figure 5.5a, unless almost the entire spectrum is deflated. For deflating less than 10% of the singular triplets—the most realistic situation—model and experimental results agree. Beyond that number, the experiment performs worse than predicted. However, the model still captures the overall effect and recommends avoiding deflation altogether. In Figure 5.5b the effect of deflation is beneficial but limited. The disagreement between model and experiment is about 10%, once again the model can be used to predict the general outcome effectively.

In summary, the presence of non-random singular vectors could generate a few discrepancies, but these and other extensive experiments show that our model is useful in predicting the overall effect of SVD deflation. Specifically, even rough knowledge of the particular singular value spectrum is enough for the model to determine whether deflation would be valuable or hurtful. Finally, we emphasize the small sizes of the above matrices.

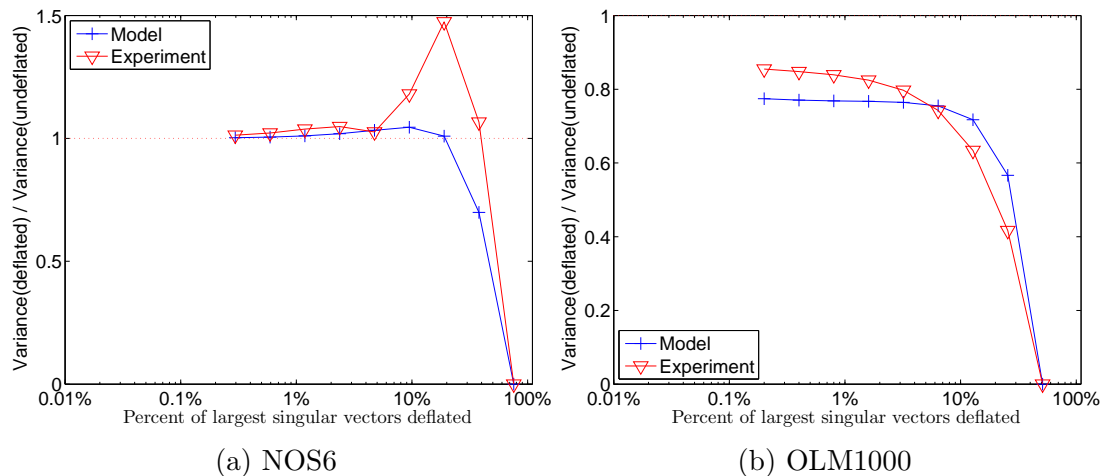


FIG. 5.5: Matrix NOS6 is symmetric, with $N = 675$ and a condition number of $7.65049e+06$. Matrix OLM1000 is non-symmetric, with $N = 1000$ and a condition number of $1.48722e+06$.

In larger applications, the singular vectors are more likely to behave like random ones.

5.3.2 Experiments with Gaussian vectors

Here, we take a brief aside from the examination of \mathbb{Z}_4 noise and study the effects of deflation under Gaussian distributed trace estimation. As stated previously in Section 5.1, $\text{Var}[\text{Tr}_{\text{Gaussian}}(A)]$ only depends on the sum of the squares of the singular values. Therefore, random matrix theory is not necessary to factorize out the effects of the singular vectors and an exact prediction of variance reduction can be made with only knowledge of the singular values.

To test this, the four matrices from the preceding subsection are deflated under the Gaussian estimator. Equations 5.10 and 5.11 are used as the model.

Figures 5.6a, 5.6b, 5.7a, and 5.7b illustrate perfect agreement between our formulas and the experimentally computed variance. Moreover, a comparison of Figures 5.5b and 5.7b conveys an example of the statement that deflation is never hurtful with Gaussian noise.

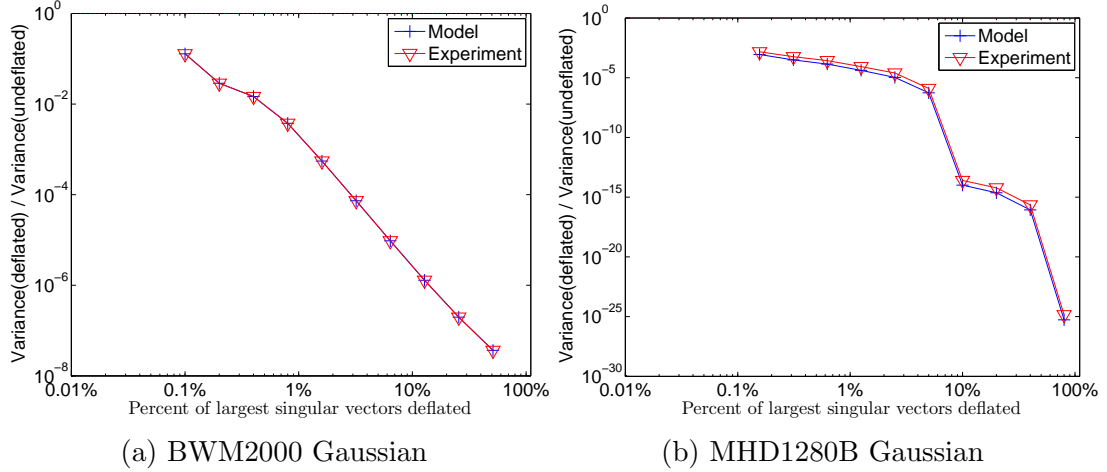


FIG. 5.6: Here the variance reduction of matrix BWM2000 and matrix MHD1280B is shown with Gaussian noise. We remind the reader that the sizes of these matrices are $N = 2000$ and $N = 1280$ respectively, with condition numbers of $2.37869e+5$ and $4.74959e+12$ correspondingly. Both matrices are real, non-symmetric.

Now all the tools are in place to investigate a comparison of the effects of SVD deflation between \mathbb{Z}_4 and Gaussian noise. From (4.29) and (4.31), it is clear that without deflation \mathbb{Z}_4 noise is optimal. Once deflation is turned on however, it is not obvious if this remains true. With the NOS6 matrix, we have already demonstrated an example in which deflation is hurtful with the \mathbb{Z}_4 estimator, while beneficial under the Gaussian one. It is worth mentioning that deflation does not appear to be useful in either case though, as a large number of low lying singular triplets must be projected out to obtain significant variance reduction.

Figures 5.8a, 5.8b, and 5.9b give a comparison of variance reduction with SVD deflation between Gaussian and \mathbb{Z}_4 noise. Since the model introduced in Section 5.1 does not exactly predict the variance for the \mathbb{Z}_4 case, we use the experimentally determined variance for the comparison. Since NOS6 is real and symmetric, Gaussian noise is compared to \mathbb{Z}_2 in Figure 5.9a rather than \mathbb{Z}_4 vectors. While this introduces an extra factor of 2 in contrast to the \mathbb{Z}_4 case, it keeps all floating points to be only real-valued, akin to Gaussian

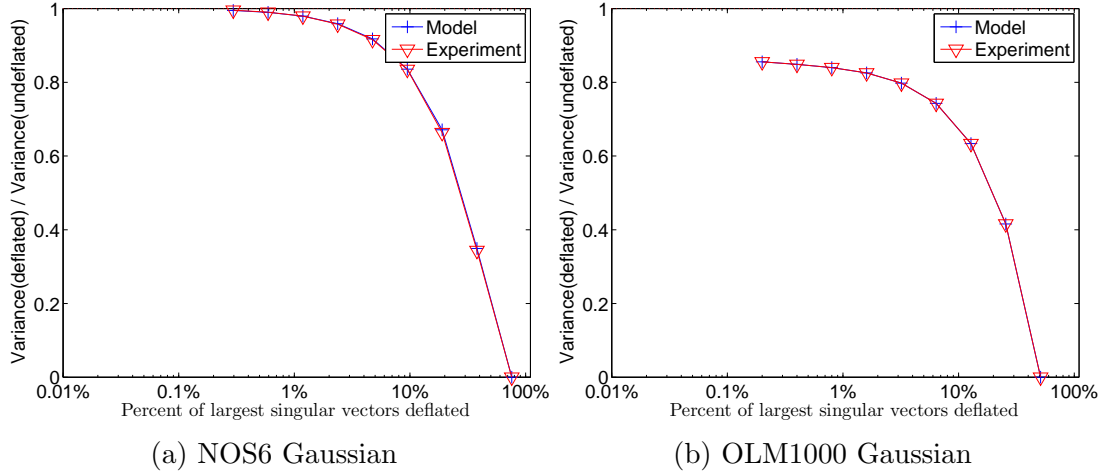


FIG. 5.7: NOS6 is $N = 675$ symmetric matrix, with and a condition number of $7.65049e+06$. OLM1000 is a $N = 1000$ non-symmetric matrix, with and a condition number of $1.48722e+06$. Both are deflated with Gaussian noise.

noise. Therefore, for real and symmetric matrices, evaluating Gaussian noise to \mathbb{Z}_2 is a fairer comparison.

For matrices BWM2000 and MHD1280B, deflation works well and the general behavior between the two cases can be seen to be similar in Figures 5.8a and 5.8b. We alert the reader of the log scale in Figure 5.8b and note that the \mathbb{Z}_4 case is a factor of 2-5 smaller than Gaussian noise throughout. In Figure 5.9a it is clear that although deflation always has a positive effect in the Gaussian case, the extra variance from starting with this noise is much too high to compete with the smaller variance from \mathbb{Z}_2 vectors. For either choice of noise however, deflation does not appear practical for the NOS6 matrix. Figure 5.9b behaves similarly to Figure 5.9a. In all cases studied here, deflation with \mathbb{Z}_4 vectors (or \mathbb{Z}_2 vectors in the case of NOS6) performs better than Gaussian noise. The possibility that Gaussian vectors behave better under deflation does not outweigh their naturally higher starting variance compared to \mathbb{Z}_4 vectors.

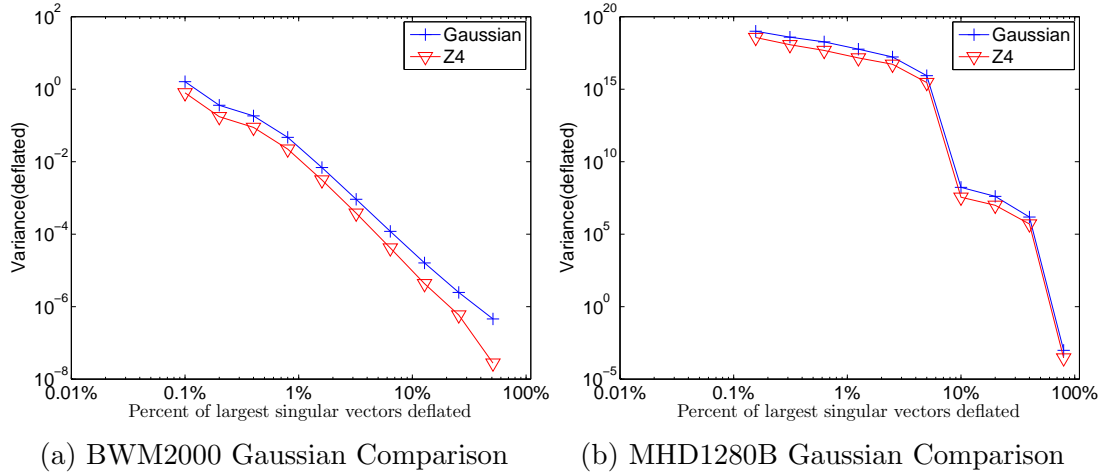


FIG. 5.8: The variance reduction of matrices BWM2000 and MHD1280B is shown with Gaussian and \mathbb{Z}_4 noise. Both matrices are real and non-symmetric.

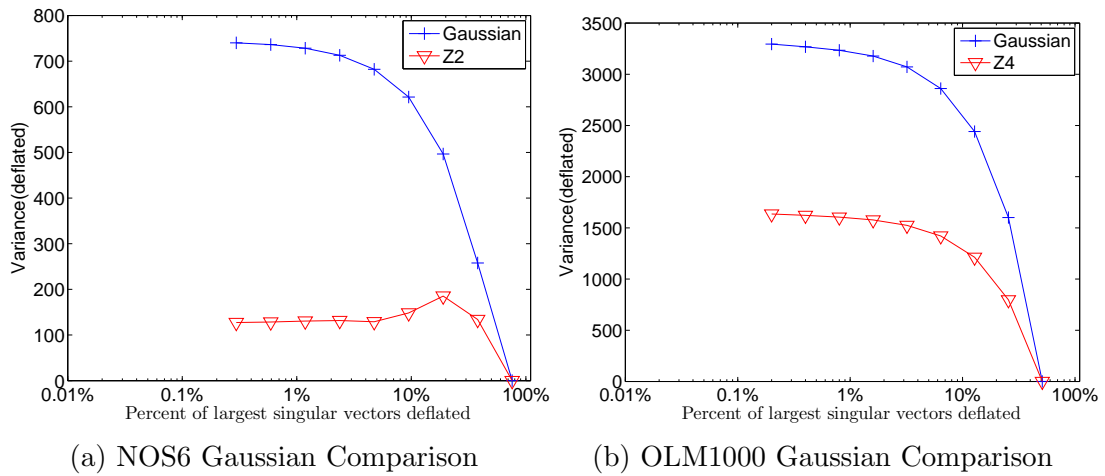


FIG. 5.9: A comparison of variance reduction for matrices NOS6 and OLM1000 is shown. Since NOS6 is real and symmetric, its comparison is done with \mathbb{Z}_2 noise.

5.4 Application to Lattice QCD

Now that the effects of SVD deflation have been studied in general, we focus on its application to the calculation of disconnected diagrams in Lattice QCD. This discussion is split into two parts. The first subsection describes the specific strategy we employed to compute low singular triplets of the Dirac matrix. We take advantage of state-of-the-art software packages and algorithms to enable our eigensolver to compute one thousand low eigenpairs, one of the largest eigenvalue calculations performed to date in Lattice QCD. The second subsection outlines an advanced trace estimation algorithm based on SVD deflation and HP.

5.4.1 How to obtain the deflation space

We are interested in the trace of the inverse of a matrix, so we need to compute its smallest singular triplets. Then, we apply the Hutchinson method on the deflated matrix by solving a series of linear systems of equations. It would have been desirable to compute the deflation space from the search spaces built by the iterative methods for solving these linear systems. This idea has been explored effectively for Lattice QCD in the past [58, 59, 60]. However, such methods are not suitable for our current problem for the following reasons.

First, methods such as GMRESDR or eigBICG produce approximations to the lowest magnitude eigenvalues of the non Hermitian matrix A . Much experimentation has shown that this eigenspace is not optimal in reducing the variance of the Hutchinson method. To produce the smallest singular triplets we would have to work with eigCG on the normal equations $A^\dagger A$ [59]. Second, only the lowest few eigenpairs produced by eigCG are accurate. The rest may have a positive effect on speeding up the linear solver, but they do not seem adequate for variance reduction. Third, and most important, we are interested in

large scale problems for which unpreconditioned eigCG would not converge in a reasonable amount of time. However, if a preconditioner M^{-1} is used, all the above methods find the eigenpairs of $M^{-1}A$ or of $M^{-1 \dagger}M^{-1}A^{\dagger}A$. These may help speed up the linear solver but are not relevant for deflating A^{-1} for variance reduction.

The alternative is to compute the deflation space through an explicit eigensolver on $A^{\dagger}A$. This is a challenging problem for our large problem sizes because the lower part of the spectrum becomes increasingly dense and eigenvalue methods converge slowly. Although Lanczos type methods are good for approximating large parts of the spectrum, they cannot use preconditioning so they are unsuitable for our problems.

For our computations of disconnected diagrams, we have used the state-of-the-art library PRIMME (PReconditioned Iterative MultiMethod Eigensolver) [61] which offers a suite of near-optimal methods for Hermitian eigenvalue problems. Among several unique features, PRIMME has recently added support for solving large scale SVD problems, including preconditioning capability, something that is not directly supported by other current software. We remind the reader from Section 4.2.3 of a multi-group, multi-year effort that has resulted in a highly efficient preconditioner which is based on domain decomposition and adaptive Algebraic Multigrid (AMG) [46]. AMG is a game changer, but it has only been used to solve linear systems of equations. We employ AMG as a preconditioner in PRIMME to find 1000 lowest singular triplets. For most methods in PRIMME, AMG accelerates the number of iterations by orders of magnitude and results in wallclock speedups of around 30.

To obtain the best performance for our specific problem in Lattice QCD, we experimented with various PRIMME methods and parameters, and AMG configurations. Our final setup of PRIMME called the `L0PBCG_Orthobasis_Window` method. Details of our optimizations can be found in [56].

The AMG software provides a solver for a non-Hermitian linear system $Ax = b$,

not just a preconditioner. There are three levels of multigrid with a GCR smoother at each level [46]. Because PRIMME needs a preconditioner for the normal equations, $A^\dagger A \delta = r$, each preconditioning application involves two calls to AMG to solve the two systems approximately, $A^\dagger y = r$ and $A \delta = y$. We found 4 GCR iterations at the fine level and 5 GCR iterations at each of two coarse levels to be optimal. Preconditioning for eigenvalue problems differs from linear systems in the sense that it should approximate $(A^\dagger A - \sigma I)^{-1}$ to improve eigenvalues near σ . In our AMG preconditioner σ is zero, so we expect the quality of the preconditioner to wane as the eigensolver locks on to eigenvalues inside the spectrum. However, the lowest part of the spectrum is quite clustered and such deterioration from the AMG preconditioner is small. We were able to use this setup to compute one thousand low lying singular vectors of the Dirac matrix on two different ensembles.

5.4.2 SVD deflation and HP trace algorithm

Given k eigenpairs (Λ, V) of the normal equations, the left singular vectors can be obtained as $U = AV\Sigma^{-1}$, where $\Sigma = \Lambda^{1/2}$. Following (5.2), we can decompose $\text{Tr}(A^{-1}) = \text{Tr}(A_D^{-1}) + \text{Tr}(A_R^{-1}) = \text{Tr}(V\Sigma^{-1}U^\dagger) + \text{Tr}(A^{-1} - V\Sigma^{-1}U^\dagger)$. Using the cyclic property of the trace, we have $\text{Tr}(V\Sigma^{-1}U^\dagger) = \text{Tr}(\Sigma^{-1}U^\dagger V) = \text{Tr}(\Lambda^{-1}V^\dagger A^\dagger V)$. This means that the trace of A_D^{-1} can be computed explicitly through k matrix vector multiplications and k inner products. Similarly, we see that $\text{Tr}(A_R^{-1}) = \text{Tr}(A^{-1} - V\Lambda^{-1}V^\dagger A^\dagger)$, so the quadratures required in Hutchinson's method can be computed as $z^\dagger A^{-1} z$ and $z^\dagger V\Lambda^{-1}V^\dagger(A^\dagger z)$. This means that we can avoid the significant storage of U .

We now have all the components to run the deflated Hutchinson method. Moreover, the same setup with deflation can be used with HP vectors. HP uses an implicit distance- d coloring of the lattice to pick the probing vectors as certain permutations of Hadamard

vectors that remove all trace error that corresponds to A_{ij}^{-1} elements with i, j having up to d Manhattan distance in the lattice. The hope is that deflation removes error in a complementary way from HP and the two techniques together lead to faster convergence. Algorithm 6 summarizes our approach.

Algorithm 6 $Trace = \text{deflatedHP}(A)$

```

1:  $[\Lambda, V] = \text{PRIMME}(A^\dagger A)$ 
2:  $T_D = \text{Tr}(\Lambda^{-1} V^\dagger A^\dagger V)$ ;  $T_R = 0$ 
3:  $z_0 = \text{randi}([0, 3], N, 1)$ ;  $z_0 = \text{exp}(z_0 \pi i / 2)$ 
4: for  $j = 1 : s$  do
5:    $z_h = \text{next vector from Hierarchical Probing or other scheme}$ 
6:    $z = z_0 \odot z_h$ 
7:   Solve  $Ay = z$ 
8:    $T_R = T_R + z^\dagger y - z^\dagger V \Lambda^{-1} V^\dagger (A^\dagger z)$ 
9:    $Trace = T_R / j + T_D$ 
10: end for

```

In Lattice QCD, we are often interested in computing $\text{Tr}(\Gamma A^{-1})$ for several different Γ matrices whose application to a vector are inexpensive to apply. In such cases, the SVD decomposition (5.2) still applies, $\text{Tr}(\Gamma A^{-1}) = \text{Tr}(\Gamma V \Lambda^{-1} V^\dagger A^\dagger) + \text{Tr}(\Gamma A^{-1} - \Gamma V \Lambda^{-1} V^\dagger A^\dagger)$. The computations are similar to Algorithm 6, with a Γ matrix vector product inserted at each step. Therefore, the computational cost of the SVD subspace can be amortized by reusing the deflation space to compute traces with multiple Γ matrices.

5.5 Numerical experiments in Lattice QCD

In this section, we present numerical results from experiments with two representative Dirac matrices. Both are from $32^3 \times 64$, $\beta = 6.3$ Clover improved Wilson ensembles. In both cases, the pion mass was about 300 MeV. However, the first matrix comes from an ensemble with 3 flavors of dynamical quarks, whose masses were turned to match the physical strange quark mass. In this case we employed a lower quark mass (quark mass

$m_q = -0.250$ in lattice units) for our numerical experiments in order to achieve a more singular matrix. In the second case, the ensemble from which we selected the Dirac matrix is one with 2 light quark flavors and one strange quark. The strange quark is again, at its physical value, and the light quarks have masses -0.239 that result in 300 MeV pions. The interested reader can find further details about these ensembles in [62, 63]. Subsequently, we will refer to the matrix with a quark mass of $m_q = -0.250$ as the Dirac operator from ensemble A, and the $m_q = -0.239$ mass matrix as the Dirac matrix from ensemble B.

The above matrices have a size of $N = 25,165,824$ and condition numbers of 1747 and 1788 respectively. As discussed earlier, the subspaces were obtained using PRIMME set to the `LOBPCG_Orthobasis_Window` method with a tolerance of 10^{-2} and a block size of 30 [61]. This was supplemented with a three level AMG preconditioner with 4^4 and 2^4 blocking and a fine/coarse maximum iteration count of 4 and 5 respectively [46].

5.5.1 Synergy between deflation and hierarchical probing

Depending on the conditioning of the matrix, the HP method has enjoyed improvements of as much as an order of magnitude compared to the basic Monte Carlo estimator [55]. In Figures 5.10a and 5.10b we present results of Algorithm 6 with the ensemble A and ensemble B matrices respectively, where HP is augmented by deflation. The error bars on the variance were estimated with the Jackknife resampling procedure on 40 runs of Algorithm 6 with different z_0 noise vectors. Local minima appear on the y axis of both plots at every power of two. This is a characteristic of the HP method, which relies on cancellations of the nested coloring algorithm to probe higher distances [55]. At least one order of magnitude improvement in variance is observed with deflation over HP alone.

Additionally, we compute the speedup of HP and deflated HP compared to the basic

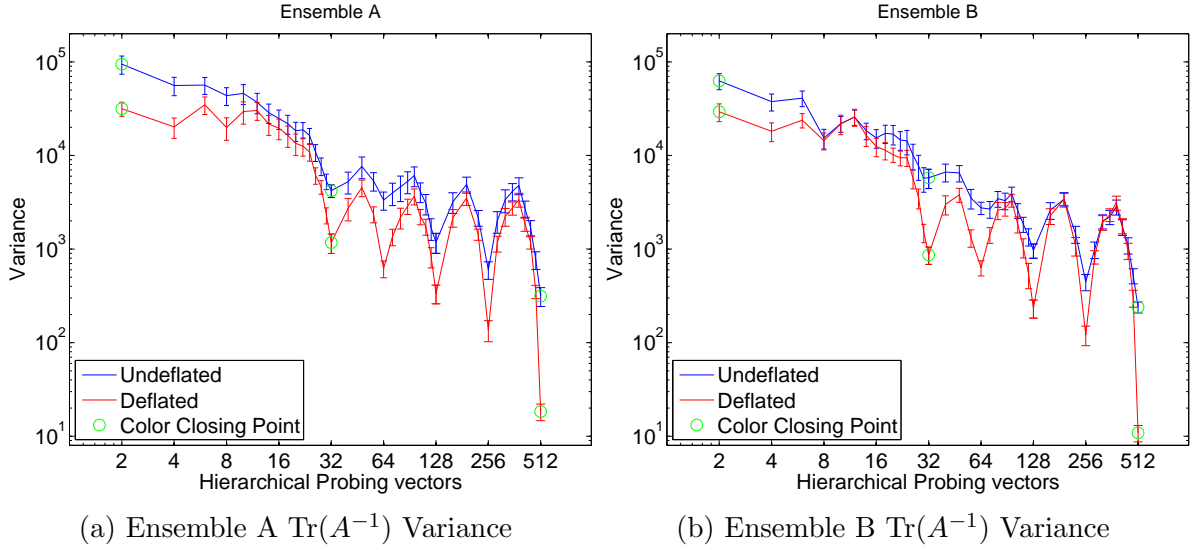


FIG. 5.10: Above is the variance of the hierarchical probing trace estimator with and without deflation. The full 1000 vector subspace is used as the deflated operator in red. Complete color closings are marked with green circles. For the ensemble A matrix, a factor of 15 is achieved in variance reduction between deflated and undeflated probing. Deflation yields over a factor of 20 reduction of variance for the ensemble B matrix.

MC estimator as

$$R_s = \frac{V_{stoc}}{V_{hp}(s) \times s}.$$

Here, V_{stoc} is the variance from the pure noise MC estimator, and $V_{hp}(s)$ is the HP variance computed with Jackknife resampling over the 40 runs. The factor of s is the number of probing vectors, and it is used to normalize the speedup ratio since the error from random noise scales as $(\frac{V_{stoc}}{s})^{1/2}$. The speedup for both ensembles are displayed in figures 5.11a and 5.11b. HP alone yields speedups of 2-3 in contrast to the factor 10 in [55] on the same matrix from Ensemble B. The difference is that in [55], the quark parameter was set to the strange quark mass while here it is set to the light quark mass, which yields the much more ill conditioned matrix in Figure 5.11b. Deflation and HP together, however, achieve a factor of 60 speedup over the original Monte Carlo method. We elaborate on this further below.

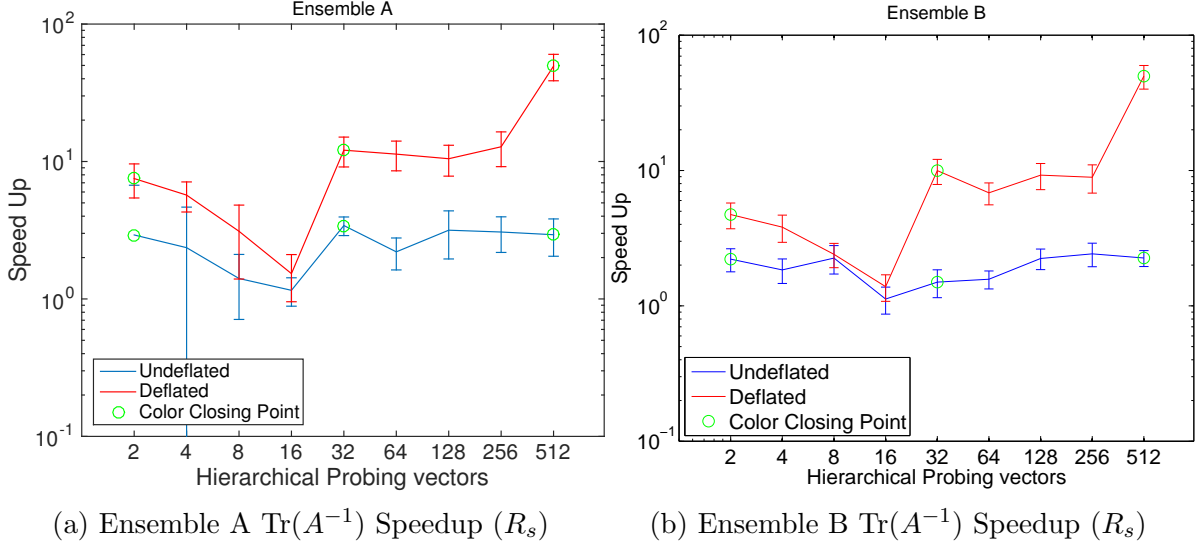


FIG. 5.11: Speedup of the combined deflated HP estimator compared to pure \mathbb{Z}_4 noise is shown. The speedup to basic MC is estimated for both HP alone and HP with deflation. The errors are computed with Jackknife resampling.

It is apparent that deflation and HP synergize to produce an extremely effective trace estimator. The idea of HP is based on the local decay of the Green's function. By assuming that the neighbors of a source node in matrix A will have weights in A^{-1} that decay with their distance from the source, HP kills the error from progressively larger distance neighborhoods. This works well for well conditioned matrices, but for ill conditioned ones A^{-1} is dominated by contributions of its near null space. Such contributions are typically non-local, which are not captured by HP. Deflation, however, captures these contributions exactly and removes them, resulting in a much easier structure for HP to work on. In Lattice QCD, this synergy completely resolves the scaling problem as the mass approaches the critical mass, and significantly reduces the effects of lattice size.

We investigate this synergy experimentally on the matrix from ensemble A. We seek to quantify the remaining variance on the original matrix ($\|\tilde{A}^{-1}\|_F^2$), after applying deflation ($\|\tilde{A}_R^{-1}\|_F^2$), after applying 32 HP probing vectors H ($\|(HH^H) \odot A^{-1}\|_F^2$), and after applying both deflation and HP ($\|(HH^H) \odot A_R^{-1}\|_F^2$). Let B denote any of these four matrices. Since

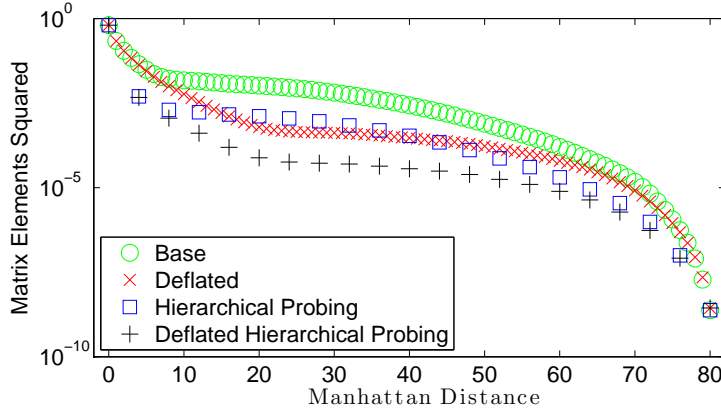


FIG. 5.12: The sum of squared absolute values of matrix elements at specified Manhattan distances from the corresponding diagonal elements for 10 randomly sampled rows. Base case is the original Monte Carlo method. Deflation refers to the Monte Carlo with deflation. HP and deflated HP refer to a space spanned by the 32 hierarchical probing vectors. A combination of HP and deflation suppresses the sum of matrix elements by orders of magnitude more than probing or deflation alone.

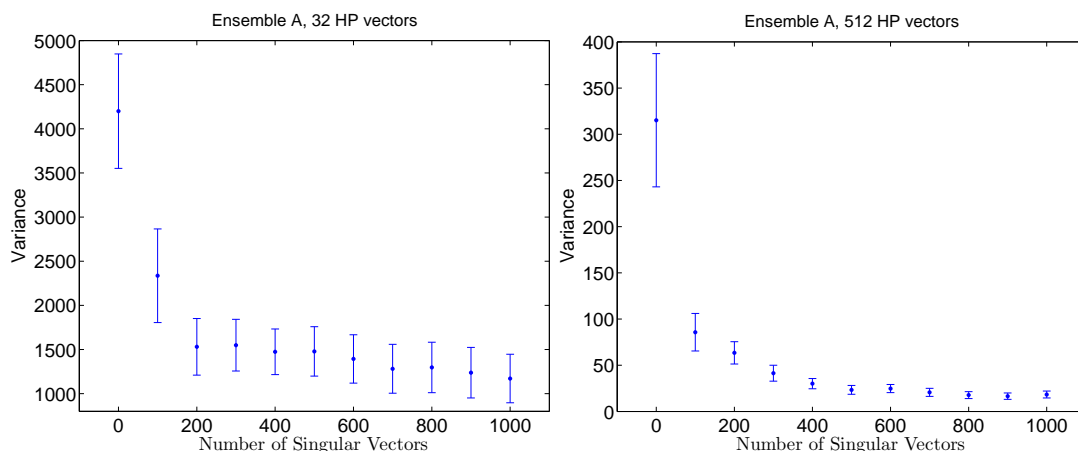
we cannot compute $\|B\|_F$ explicitly, we randomly sample 10 of its rows, denoting this set as S . Then for each corresponding lattice node $i \in S$, we find all its m_d neighbors j that are d hops away in the lattice (i.e., its Manhattan distance- d neighborhood) and sum their squared absolute values $|B_{ij}|^2$. Averaging these over all m_d neighbors and all nodes in S gives us an estimate of how much variance remains from elements at distance d . These W_d are plotted in Figure 5.12,

$$W_d = 1/|S| \sum_{i \in S} \sum_{j \in \mathcal{N}_d} |B_{ij}|^2 / m_d, \quad \text{where } \mathcal{N}_d = \{j : \text{dist}(i, j) = d\} \text{ and } m_d = |\mathcal{N}_d|.$$

The figure shows how HP eliminates the variance from the first 3 distances and repeats this pattern in multiples of 4 (1,2,3,5,6,7,...) [55]. While probing eliminates better short-distance variance, deflation is better at long-distance. Combining them achieves a much greater reduction in variance than either of the two alone.

5.5.2 Varying the SVD deflation space

We also study the effect of the size of the deflation SVD subspace. By saving all inner products performed in the trace estimator, we are able to play back the trace simulation deflating with different numbers of singular triplets. We combine deflation and HP and report results for 32 and 512 probing vectors, which represent the proper color closings for HP on a 4D lattice [55]. As before, the error bars are obtained from 40 different runs of Algorithm 6 with different z_0 .

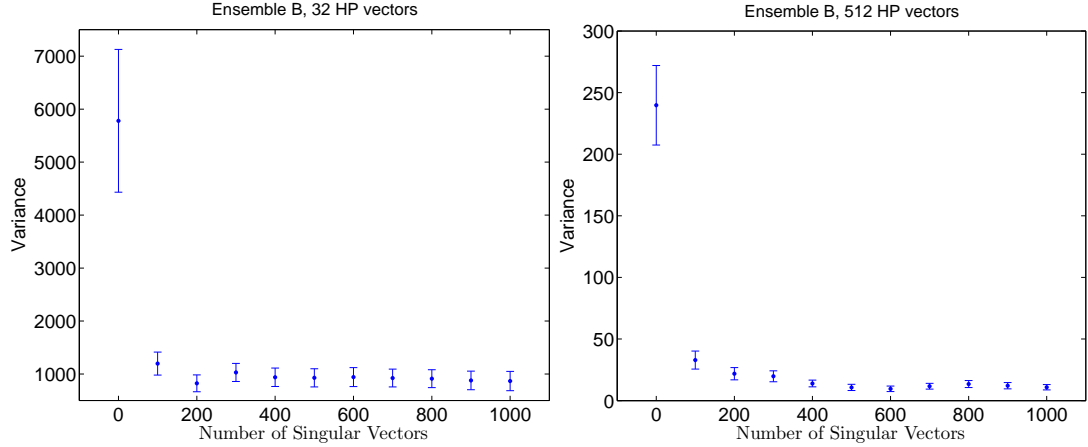


(a) Ensemble A SVD and 32 HP vectors (b) Ensemble A SVD and 512 HP vectors

FIG. 5.13: Variance for the ensemble A matrix as a function of the deflated SVD subspace dimension at two color closing points of HP. The left plot is with 32 probing vectors, the right is with the full 512.

Figure 5.13a shows that deflation with 200 singular vectors reduces variance by a factor of 3, and beyond 200 little improvement is gained. In Figure 5.13b, HP has removed the error for larger distances and therefore it can use more singular vectors effectively, yielding more than an order of magnitude improvement. Still there is potential for computational savings since 500 singular vectors have the same effect as 1000. Figures 5.14a and 5.14b display similar attributes for the ensemble B matrix.

These experiments illustrate that the optimal number of vectors to be used in each



(a) Ensemble B SVD and 32 HP vectors (b) Ensemble B SVD and 512 HP vectors

FIG. 5.14: Variance for the matrix from ensemble B, as a function of the deflated SVD subspace dimension at two color closing points of HP. The left plot is with 32 probing vectors, the right is with the full 512.

of the two techniques depends on each other. This is only an issue if one needs to figure out how many singular vectors to compute a priori, because if these are already available, their application in the method is not computationally expensive. Moreover, while using a sufficiently large number of probing vectors is important, the performance of deflation seems to be much less sensitive to the number of singular vectors. Once the near null space has been removed, there are diminishing returns to deflate with bigger subspaces.

5.5.3 Wallclock timings and efficiency

Implementing HP with deflation requires an additional setup cost from finding the SVD space. In Lattice QCD, this cost is of little importance since the subspace may be stored and reused several times for computing various correlation functions.

Below, we show that deflation is valuable even as a “one shot method” for our Lattice QCD matrices. We investigate the case in which $\text{Tr}(A^{-1})$ only needs to be computed once, and report the time to compute the SVD, first separately and then as the overhead of the

preprocessing of Hutchinson’s method with HP. Our experiments were performed on the Cray Edison using 32 12-core Intel Ivy Bridge nodes clocked at 2.4 GHz, each with only 8 cores enabled due to memory and node topology considerations.

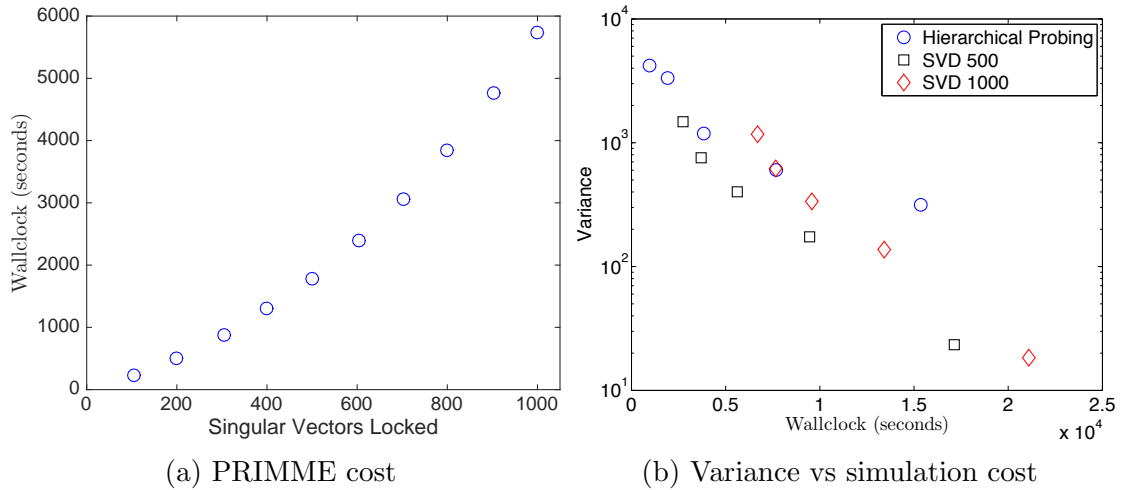


FIG. 5.15: Eigenvectors computed by PRIMME from 100 to 1000 for the matrix from ensemble A. A log plot of variance and cost. Each case displays 5 points, which represent the variance and wallclock at probing vectors 32, 64, 128, 256, and 512.

Figure 5.15a shows the timings for PRIMME as a function of the number of eigenvectors found. As more eigenvectors converge, orthogonalization costs increase, resulting in time increasing super linearly. The expected reduction in the efficiency of the AMG preconditioner as the eigensolver moves to the interior of the spectrum is in fact negligible. Obtaining 1000 eigenvectors takes 1.5 hours, while 500 vectors are computed in less than half an hour. Indeed, with the aid of the AMG preconditioner, PRIMME was able to solve for the eigenvalues of $A^\dagger A$ at a fraction of the cost of the probing estimator.

We now add the time to compute the singular space, as well as the time to perform the projections with that space, to the timings for the remaining steps of Algorithm 6. We consider two simulations; one with a deflation space of 500 vectors and one with 1000. From Figures 5.14a and 5.14bm we do not expect gain beyond 500 singular triplets. For each closing point of HP (32, 64, 128, 256, and 512 probing vectors), Figure 5.15b plots

the achieved variance as a function of total wallclock time. We observe that the variance with 500 deflation vectors at probing vector 128 is comparable to the variance of the plain HP method at 512 probing vectors. This translates to a 3-fold reduction in wallclock, even with the SVD computations included. Furthermore, at 512 probing and 500 deflation vectors, we see a 15-fold reduction in variance with the SVD time being less than 10% of total wallclock. This suggests that SVD deflation can be used equally well as a one shot method for variance reduction.

This concludes our investigation of the effects of SVD deflation. We have studied the variance reduction from deflating the lowest singular triplets in the general case and applied to disconnected quark loops in Lattice QCD. Our examination has focused primarily on \mathbb{Z}_4 noise, which we have found to be optimal both with and without deflation, although the Gaussian case was also studied. Our formulas and experiments show that the faster the singular values rise, the higher variance reduction can be expected. Furthermore, we uncovered the astonishing property that non-Hermitian matrices behave better under deflation than Hermitian ones. Both theory and real world examples were introduced with \mathbb{Z}_4 vectors that demonstrate deflation always has an affirmative effect for non-Hermitian matrices, while it may sometimes be detrimental for Hermitian ones. Conversely, for Gaussian vectors, deflation always yields a positive effect in reducing variance.

In the second half of this chapter, we concentrated on a strategy to compute disconnected correlation functions in Lattice QCD. This is equivalent to finding $\text{Tr}(\Gamma A^{-1})$, where A is the Dirac operator and Γ is one of sixteen possible matrices that represent observables such as axial, scalar, pseudoscalar, tensor, or vector matrix elements. We outlined a two-stage algorithm which first, took advantage of PRIMME and AMG to efficiently compute numerous low singular triplets of the Dirac matrix and second, removed the space of these singular values out of an HP trace estimator. We noted the synergy of HP and SVD deflation and sampled columns of our Dirac matrix to understand the reason for this feature.

We observed a factor of 15-20 improvement comparing deflated HP to HP alone, and over a factor of 60 speedup when compared to the most basic form of the Hutchinson method. Finally, the effect of varying the SVD space was studied and wallclock timings from Cray Edison were reported. With an optimal choice of parameters, over an order of magnitude of variance reduction was achieved with less than 10% additional wallclock time. Much of the content in this chapter has been referenced from [56]. In the final chapter, these methods will be applied to compute various zero-momentum transfer matrix elements in Lattice QCD.

CHAPTER 6

Selected Results

Now all the ingredients are in place to compute hadron structure observables. Armed with algorithms and techniques from the previous chapters, we present calculations of various matrix elements of the nucleon and light nuclei. The discussion is started with a study of the effects of varying the number of hierarchical probing (HP) vectors to compute zero-momentum transfer matrix elements of isoscalar and strange correlators. Subsequently, results to fitting away the excited state contamination with the methods depicted in Section 3.3 are shown for flavor diagonal matrix elements of the nucleon on two ensembles with different lattice spacings and roughly equivalent pion masses. Finally, high precision data are given for the ratio of strange sigma terms in light nuclei.

Each section is introduced with a summary of the ensemble parameters relevant to the calculation. All computations are done with clover-improved Wilson fermions for both valence and sea sectors. The ensembles used for the nucleon matrix elements are with 2+1 dynamical flavors, with the strange quark tuned near its physical value. The light nuclei calculations are performed with an ensemble at the $SU(3)$ symmetric point.

6.1 Varying HP vectors

In this section we study the effects of varying the number of HP vectors when computing the axial, scalar, tensor, and vector charges. Although variance reduction of the SVD deflated HP trace estimator is examined extensively in Chapter 5, the error on the trace from each configuration is not the only noise in a hadron three-point correlation function. Even if an exact computation of the quark loop trace was possible, the Euclidean correlators would still contain errors due to noise in the Monte Carlo over the ensemble of gauge field configurations. Therefore, the goal of our trace estimation should be in computing the trace accurately enough so that gauge noise dominates in the statistical error of the hadron matrix element of interest.

Since various observables may fluctuate differently as a function of the gauge fields, the optimal point will be dependent on which quark bilinear is of interest. We remind the reader that Algorithm 6 allows all 16 γ insertions to be computed with the same HP and SVD vectors. Thus the matrix element that is most sensitive to the accuracy of the trace per configuration will dictate the total computational resources that should be put into a quark loop calculation. In the successive subsections we present a study of the ensemble average of axial, scalar, tensor, and vector charges as a function of HP vectors for the isoscalar and strange channels.

6.1.1 Ensemble and correlator overview

Here, we introduce details of the ensemble and correlators used in the ensuing discussion. The gauge configurations were generated by the Jefferson Lab Lattice Group with 2+1 flavor clover Wilson fermions. The Monte Carlo updates were achieved with the rational hybrid Monte Carlo algorithm [64]. A single iteration of stout smearing with a weight of $\rho = .125$ was employed for the staples in the fermion action. The lattice spacing

was determined with the Wilson-flow scale, w_0 [65]. The lattice spacing is $a = .094$ fm and the pion mass is $m_\pi = 280$ MeV. The strange quark mass was tuned by fixing a ratio of hadron masses that is insensitive to the light quarks at leading order to its physical value, $\frac{2m_{K^+}^2 - m_{\pi^0}^2}{m_{\Omega^-}^2} = 0.1678$. The light quark masses were then dialed down while the strange quark mass remained fixed. The tests in Chapter 5 labeled as “Ensemble B” come from a randomly selected configuration from this ensemble. Further details of this ensemble can be found in the next section from Table 6.1.

The nucleon two-point functions were computed by the Nucleon Matrix Elements (NME) collaboration with the prescription described in Section 3.1.1. The quark propagators were computed by inverting the clover Wilson Dirac matrix on Gaussian smeared sources. The source and sink of the nucleon correlators were smeared by several applications of the three-dimensional Laplacian operator with a smearing width of $\sigma = 9$. The all mode averaging (AMA) method was used to speed up the linear solves [66, 67]. Three time-sources separated by 24 lattice units were taken with 32 low precision-sources per time-slice. One high-precision measurement per time-slice was utilized to correct the bias from the low-precision measurements. Additional details of the nucleon two-point generation may be found in [62, 63].

The three-point functions are built by multiplying the two-point correlators by the disconnected quark loops on appropriate timeslices. The forward propagating three-point correlators are defined as

$$c_f^{3pt}(t_{\text{sep}}, \tau) = c_f^{2pt}(t_{\text{sep}})L(\tau + t_0), \quad (6.1)$$

where t_{sep} is the source-sink separation, defined as $t_{\text{sep}} = t - t_0$, τ is the loop insertion time, c_f^{2pt} is the forward nucleon two-point function, and L is the quark loop. Similarly, the backward propagating three-point correlator is

$$c_b^{3pt}(t'_{\text{sep}}, \tau) = c_b^{2pt}(t'_{\text{sep}})L(t_0 - \tau), \quad (6.2)$$

where $t'_{\text{sep}} = t_0 - t$ and c_b^{2pt} is the backward propagating nucleon two-point function. For the scalar matrix element, the operator has a non-zero contribution from the vacuum which must be subtracted out:

$$\langle c_f^{3pt}(t_{\text{sep}}, \tau) \rangle = \langle c_f^{2pt}(t_{\text{sep}})L(\tau + t_0) \rangle - \langle c_f^{2pt}(t_{\text{sep}}) \rangle \langle L(\tau + t_0) \rangle, \quad (6.3)$$

and similarly for backward propagation, $\langle \rangle$ refers to the ensemble average. Roughly four hundred configurations were analyzed, with 256 HP vectors and 500 SVD vectors for the light quark loops and 512 HP vectors with 200 SVD vectors for the strange ones. In order to extract the relevant matrix elements, the ratio of the three-to-two-point correlators is taken, following the methodology summarized in Section 3.3. We caution to the reader however that only raw data of this ratio is plotted in this section and no fits are used. This is simply a study of how the accuracy of the quark loop affects the statistical errors of hadron matrix elements. Fits are performed in the final two sections of this chapter.

6.1.2 Axial charge

We plot the ratio of the three-point and two-point correlation functions for the axial charge in the isoscalar ($u + d$) channel.

In Figures 6.1a, 6.1b, 6.2a, and 6.2b the dependence on the statistical error of the axial charge is shown. The y axis is kept fixed for all figures. The correlator ratio is plotted as a function of $\tau - t_{\text{sep}}/2$, therefore the points closest to 0 on the x axis have the least excited state contributions. We alert the reader that we have only plotted a narrow

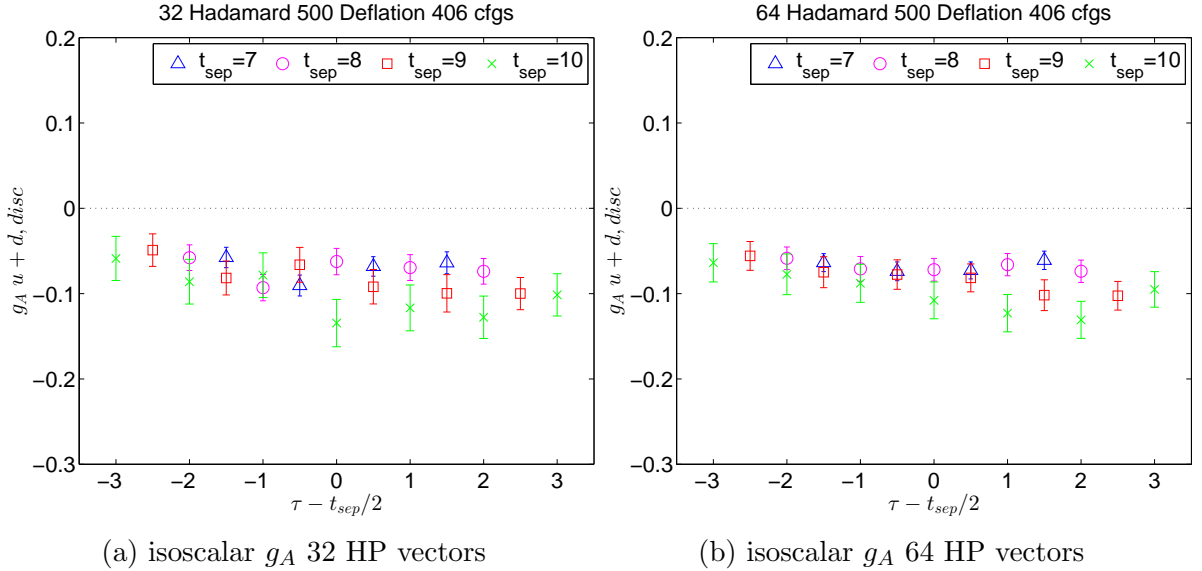


FIG. 6.1: A side by side comparison of the isoscalar disconnected axial charge at 32 and 64 HP vectors. There is a noticeable improvement scaling up to 64 HP vectors.

separation away from this zero-point and that further separations will begin to form a tail as excited state contamination grows. Noticeable reduction is observed up to 128 deflated HP vectors, however beyond this the statistical error seems to saturate, resulting in little to no improvement in 256 HP vectors.

The NME collaboration computed disconnected quark loops on 100 configurations of this ensemble using only slight advances from the basic Monte Carlo method. Their runs used 512 \mathbb{Z}_2 noise vectors improved with the hopping parameter expansion [68, 69] and the truncated solver method [66].

In Figure 6.3 a comparison between \mathbb{Z}_2 noise with slight modifications and the methods presented in this work is drawn. Even at 32 vectors, deflated HP produces smaller statistical error than 512 \mathbb{Z}_2 vectors. We note that the point $t_{\text{sep}} = 10$, $\tau = 5$ was due to smaller excited state contaminations, however shorter points in t_{sep} and τ will have less gauge noise, resulting in an even larger contrast between the two methods.

We also present data on the strange axial charge as a function of the number of

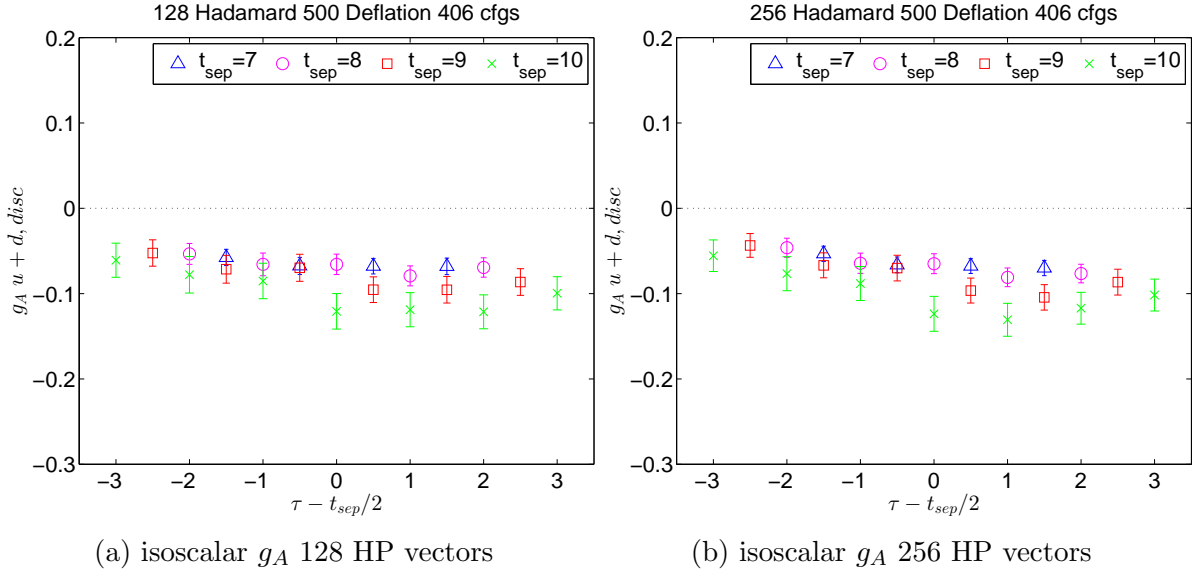


FIG. 6.2: A side by side comparison of the isoscalar disconnected axial charge at 128 and 256 HP vectors. Error in the matrix element continues to decrease with up to 128 HP vectors. No significant improvement is observed with 256 HP vectors.

deflated HP vectors. These runs are carried up to the third color closing point, or 512 HP vectors for four-dimensional probing. No comparison to Monte Carlo trace estimation is made for strange data.

Figures 6.4a, 6.4b, 6.5a, 6.5b, and 6.5c display the strange axial charge with various numbers of HP vectors. Consistent with the isoscalar case, after 128 HP vectors, error reduction is not observed.

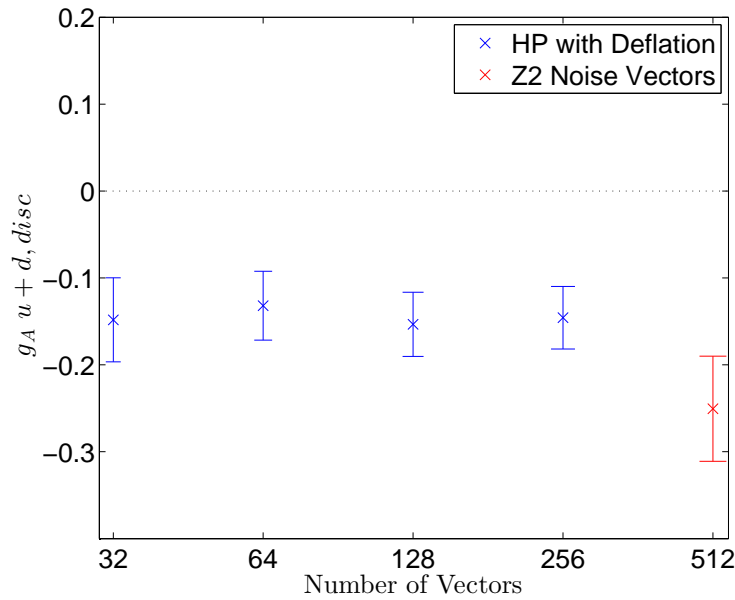


FIG. 6.3: Above, the axial charge with deflated HP is compared to a pure noise trace estimator in the isoscalar sector. The points shown are with $t_{\text{sep}} = 10$, $\tau = 5$. Both methods are performed on the same 100 configurations.

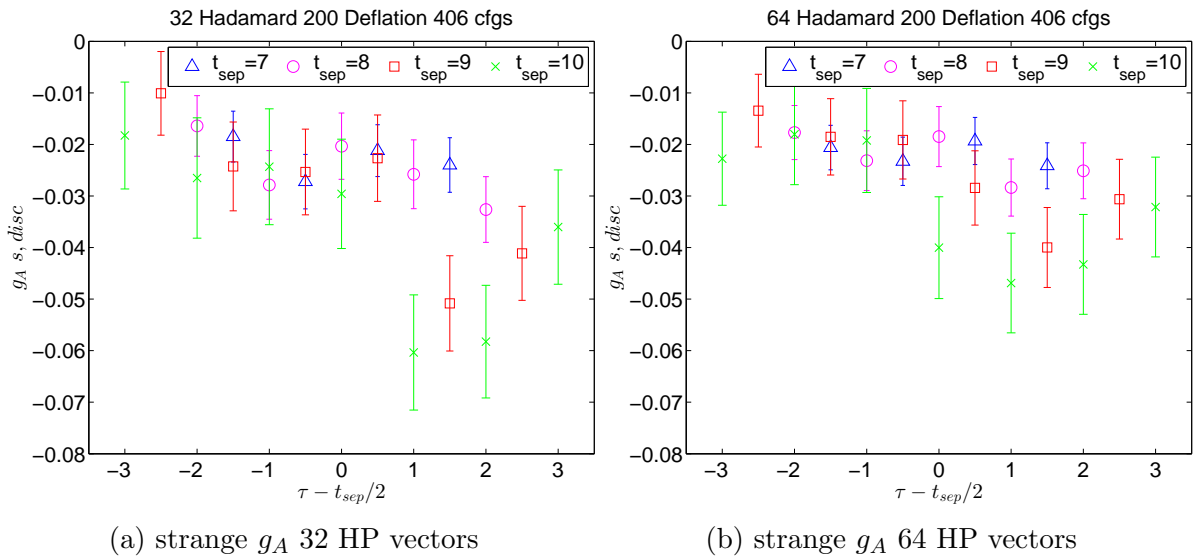
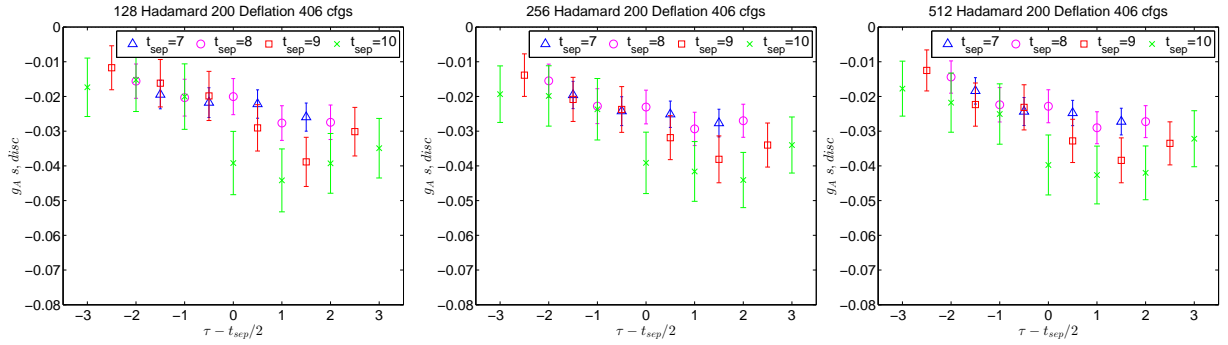


FIG. 6.4: A side by side comparison of the strange axial charge at 32 and 64 HP vectors. There is a noticeable reduction of error through 64 HP vectors.



(a) strange g_A 128 HP vectors (b) strange g_A 256 HP vectors (c) strange g_A 512 HP

FIG. 6.5: A side by side comparison of the strange axial charge at 128, 256, and 512 HP vectors is shown. Error in the data appear lower with 128 HP vectors compared to 64. No significant improvement is observed with 256 or 512 HP vectors however.

6.1.3 Scalar charge

Akin to the previous subsection, we plot the ratio of the scalar nucleon three-point function to the two-point function for the isoscalar combination. First, we show g_S varying with the number of HP vectors at a variety of different source-sink separations and loop insertions. Second, we present a comparison between deflated HP and \mathbb{Z}_2 noise runs from NME.

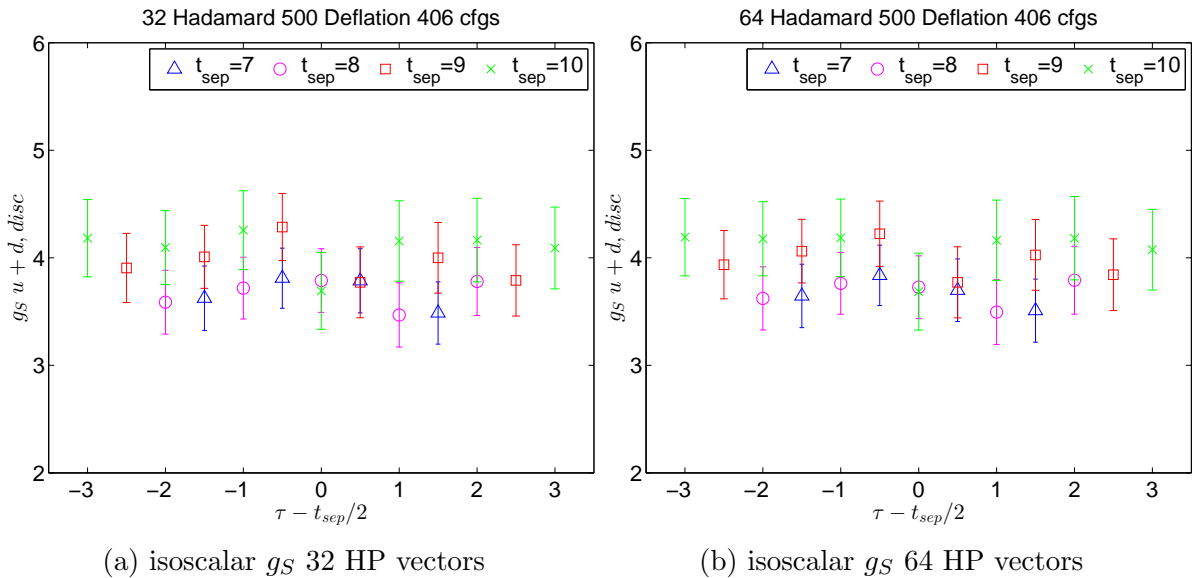


FIG. 6.6: A side by side comparison of the isoscalar disconnected scalar charge at 32 and 64 HP vectors. No difference is observed.

Figures 6.6a, 6.6b, 6.7a, 6.7b, and 6.8 show no improvement with increasing number of HP vectors or between the basic stochastic estimator and our variance reduced methods. The scalar charge is well known for fluctuating heavily in the gauge noise [70]. Hence an accurate trace on each configuration is not required. We reproduce and confirm this feature. We also show the strange scalar charge with different numbers of HP vectors.

The error in the strange scalar charge behaves identically to the isoscalar channel. We conclude that if only the zero-momentum transfer scalar matrix element is of interest, an extremely low budget trace estimator should be used.

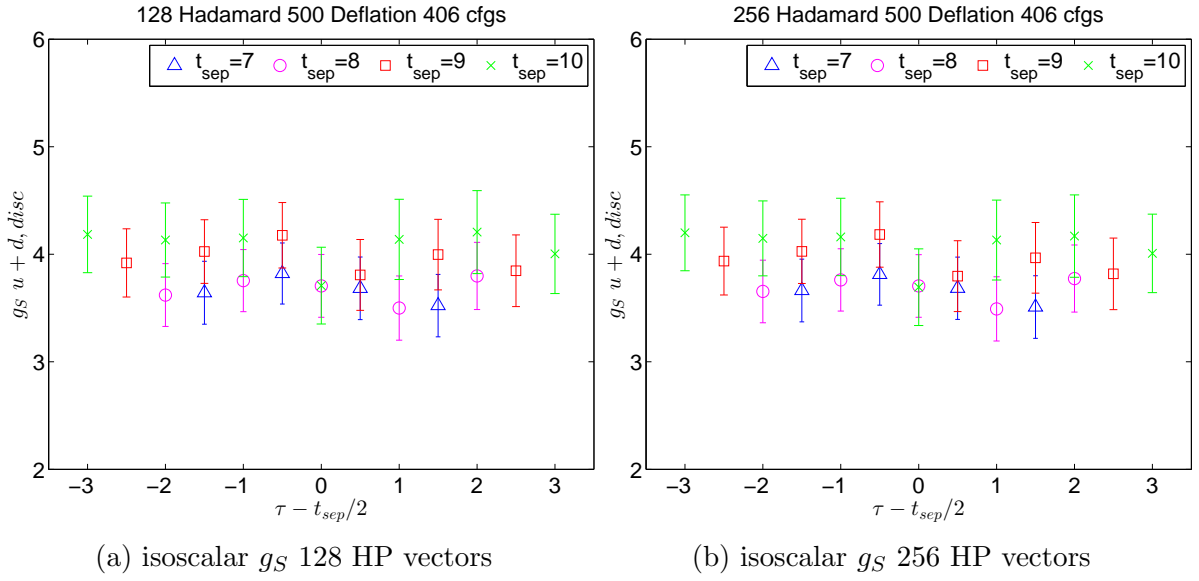


FIG. 6.7: A side by side comparison of the isoscalar disconnected scalar charge at 128 and 256 HP vectors. No difference is observed.

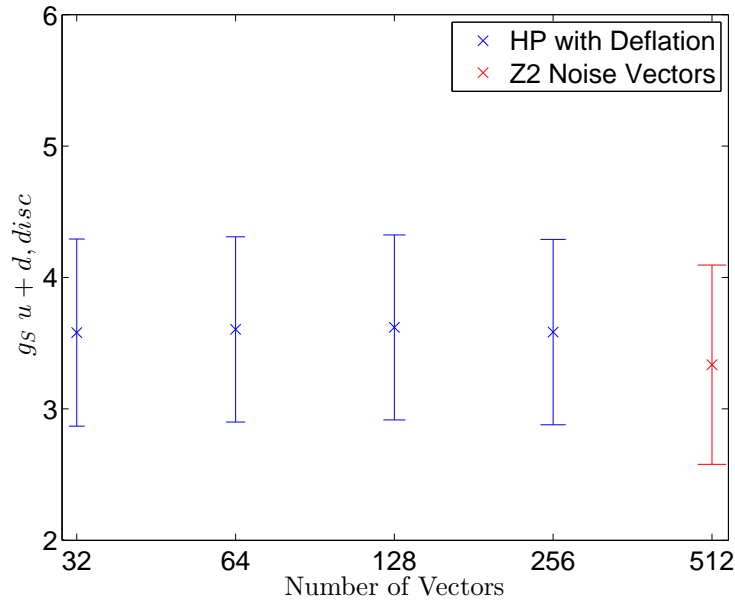


FIG. 6.8: The scalar charge with deflated HP is compared to a pure noise trace estimator in the isoscalar sector. The points shown are with $t_{sep} = 10$, $\tau = 5$. Both methods are performed on the same 100 configurations.

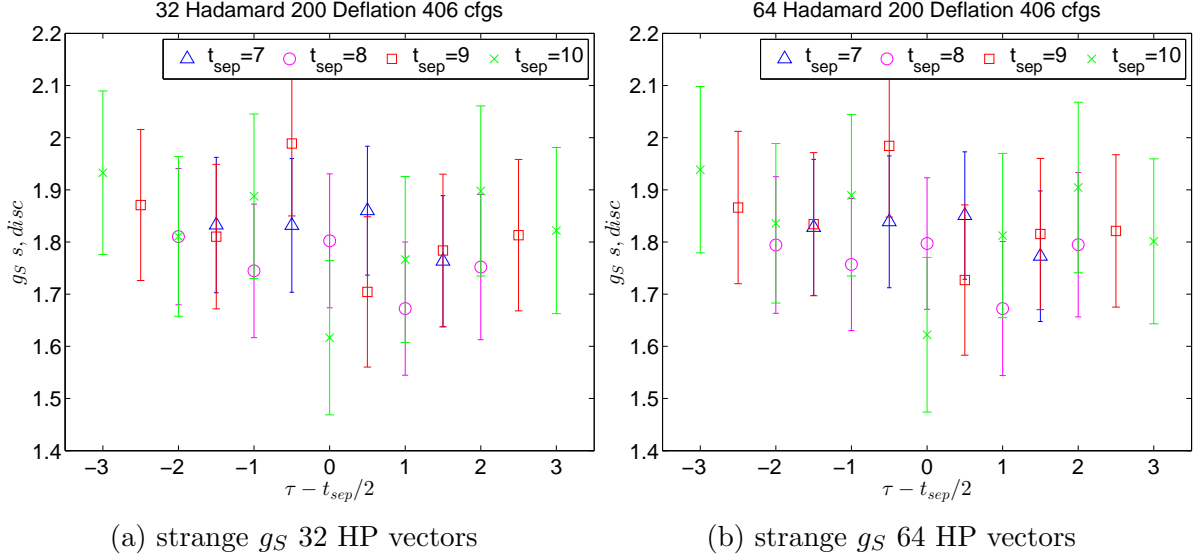


FIG. 6.9: A side by side comparison of the strange scalar charge at 32 and 64 HP vectors. No difference is observed.

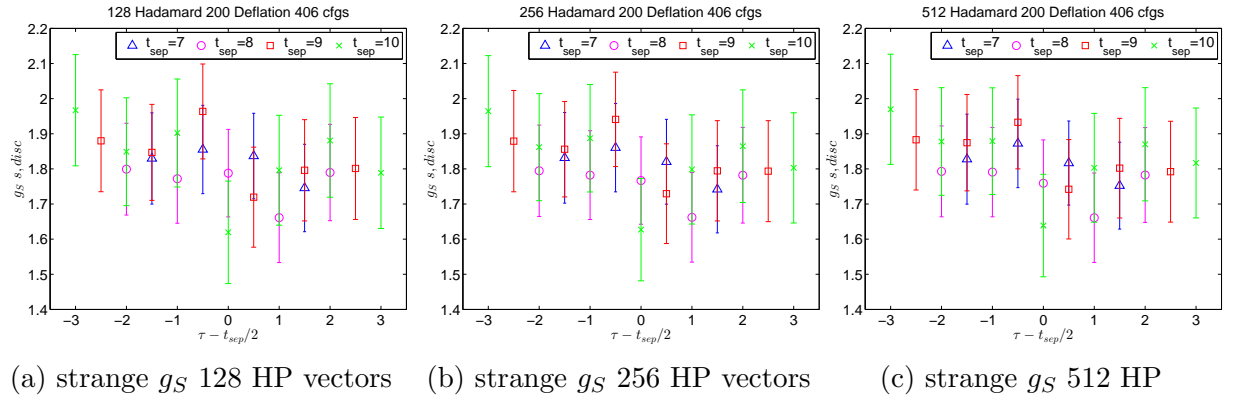


FIG. 6.10: A side by side comparison of the strange scalar charge at 128, 256, and 512 HP vectors is plotted. Statistical errors are equivalent to even the most inaccurate computation of the trace in Figure 6.9a.

6.1.4 Tensor charge

Here we show the ratio of the tensor nucleon matrix element for the isoscalar combination. A comparison to NME \mathbb{Z}_2 noise runs for a single point and full t_{sep} , τ dependence as a function of Hadamard vectors is plotted.

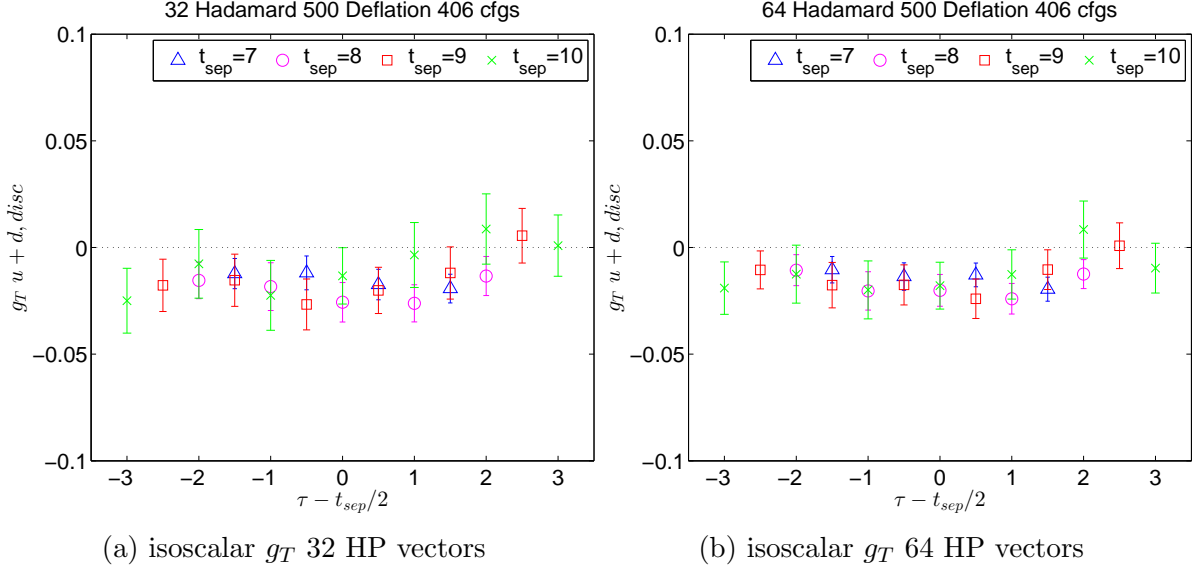


FIG. 6.11: A side by side comparison of the isoscalar disconnected tensor charge at 32 and 64 HP vectors. Error reduces from 32 to 64 HP vectors.

Figures 6.11a, 6.11b, 6.12a, 6.12b, and 6.13 display nice improvement throughout. The reduction from 64 to 128 HP vectors is particularly noteworthy. Additionally, deflated HP performs much better than \mathbb{Z}_2 noise, which gives an answer that could be consistent with zero. The strange tensor charge as a function of HP vectors is also shown.

Beyond 128 deflated HP vectors the errors in the strange tensor charge do not improve. It is possible that the strange contribution to the tensor matrix element fluctuates more from configuration to configuration. Alternatively, 200 SVD deflation vectors may reduce the variance of the strange tensor charge substantially more than 500 SVD vectors do for the light quark loops, achieving the same saturation threshold earlier in the number of HP vectors.

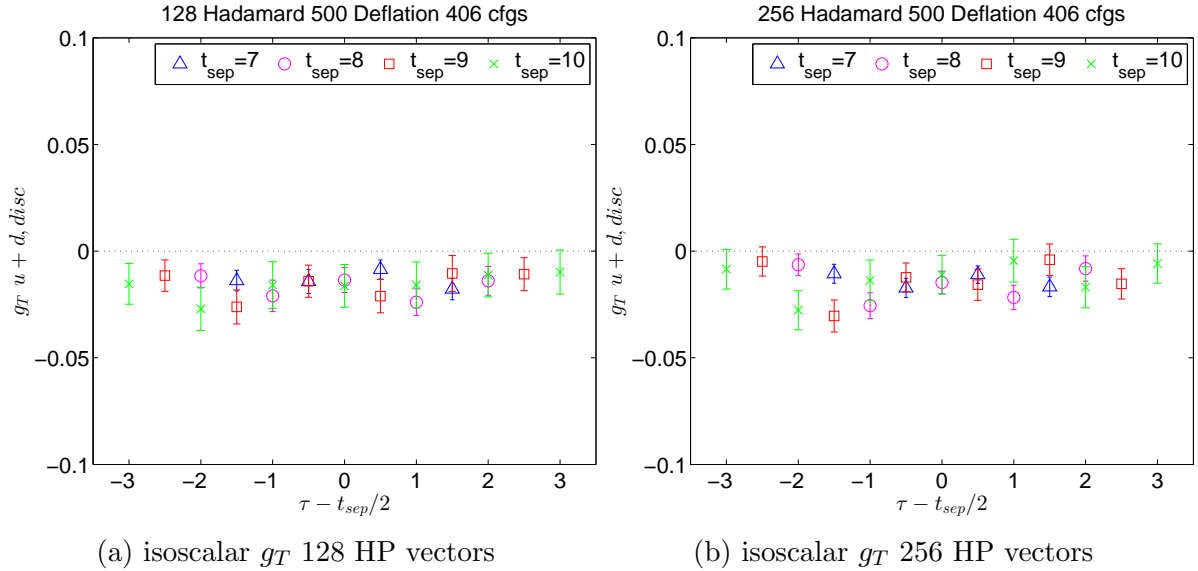


FIG. 6.12: A side by side comparison of the isoscalar disconnected tensor charge at 128 and 256 HP vectors. There is significant error reduction from 64 to 128 HP vectors. Additionally, the errors in the endpoints continue to decrease through 256 HP vectors.

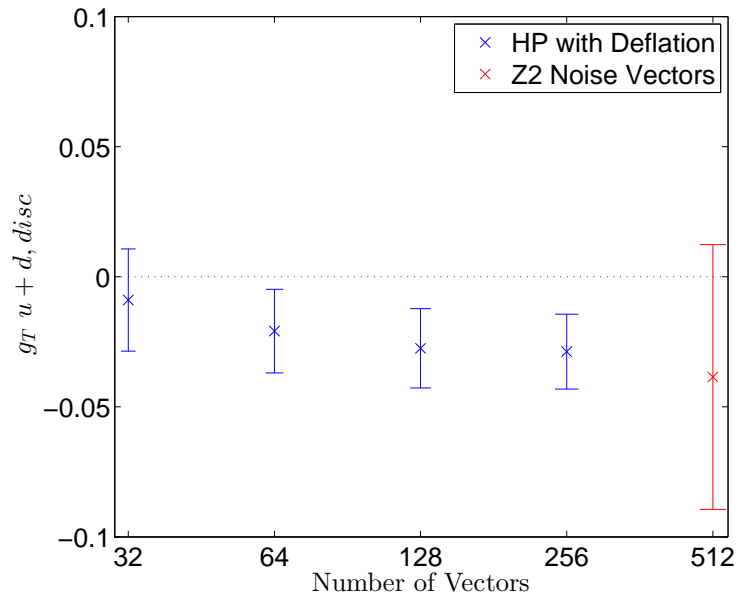


FIG. 6.13: The tensor charge with deflated HP is compared to a pure noise trace estimator in the isoscalar sector. The points shown are with $t_{sep} = 10$, $\tau = 5$. Both methods are performed on the same 100 configurations.

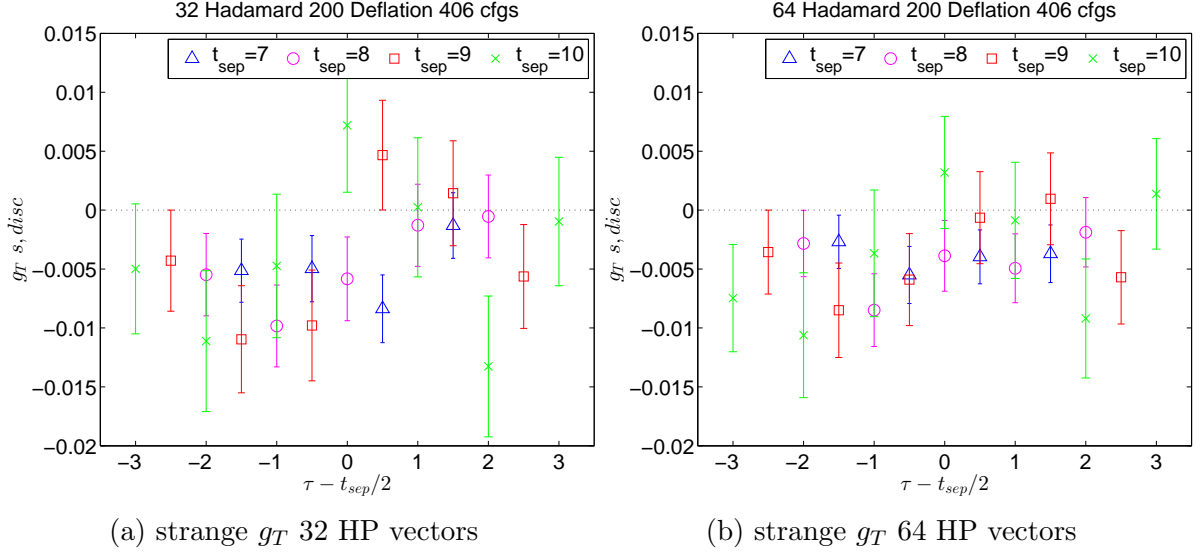


FIG. 6.14: A side by side comparison of the strange tensor charge at 32 and 64 HP vectors. Errors, especially of the endpoints, reduce at 64 HP vectors compared to 32.

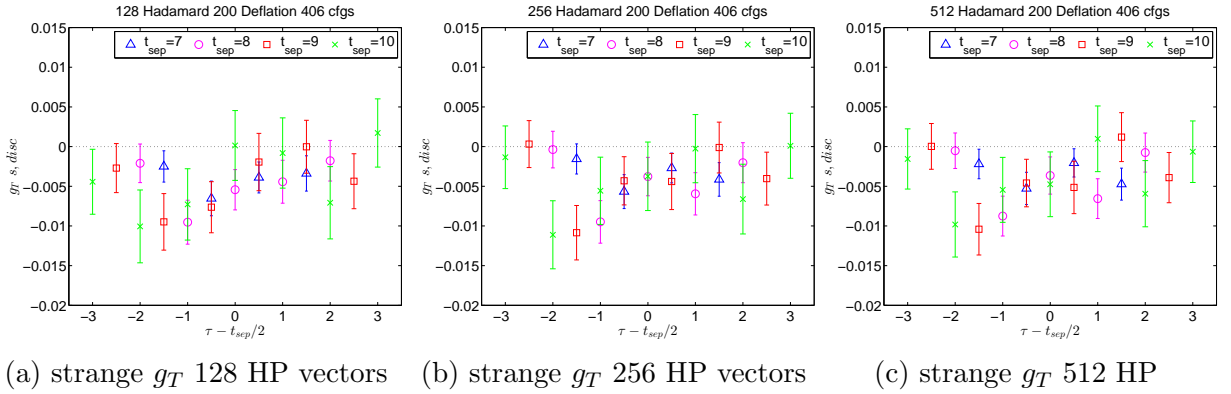


FIG. 6.15: A side by side comparison of the strange tensor charge at 128, 256, and 512 HP vectors is shown. We observe only a little improvement from 64 to 128 HP vectors for the strange contribution. No clear gain is achieved beyond 128 HP vectors.

6.1.5 Vector charge

Finally, the nucleon vector charge for the isoscalar combination is shown. A comparison to NME \mathbb{Z}_2 noise runs for a single point and full t_{sep} , τ dependence as a function of Hadamard vectors is plotted. Note that at zero-momentum transfer, the vector current simply gives the electric charge of the proton. Since this contribution comes from the connected diagrams, we expect the disconnected pieces to be zero. Nevertheless, verifying this property is a nice consistency check of the loop computations. Furthermore, gaining insight into the statistical errors of the vector charge is extremely useful for future high-precision measurements of the electric and magnetic form factors at non-zero or even high-momentum transfer [71].

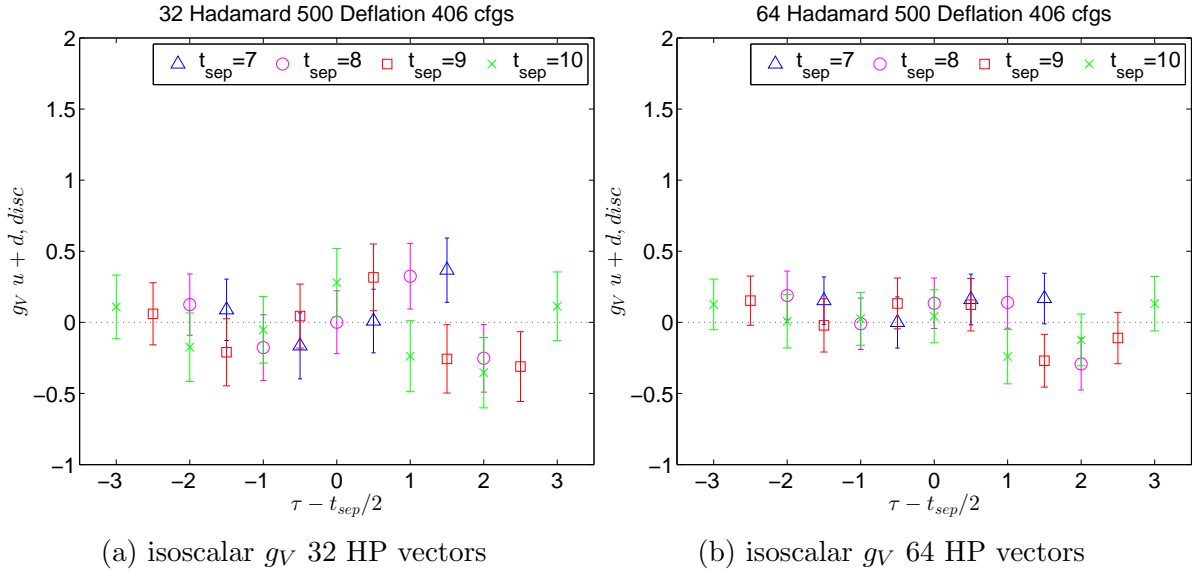


FIG. 6.16: A side by side comparison of the isoscalar disconnected vector charge at 32 and 64 HP vectors. The statistical error decreases significantly from 32 to 64 HP vectors.

Figures 6.16a, 6.16b, 6.17a, 6.17b, and 6.18 indicate a drastic drop of statistical errors as the number of HP vectors is increased. Furthermore, the data move toward zero as the trace per configuration becomes more accurate. SVD deflated HP outperforms 512 \mathbb{Z}_2 noise vectors at all points beyond 32. Nevertheless, we point out that going all the way

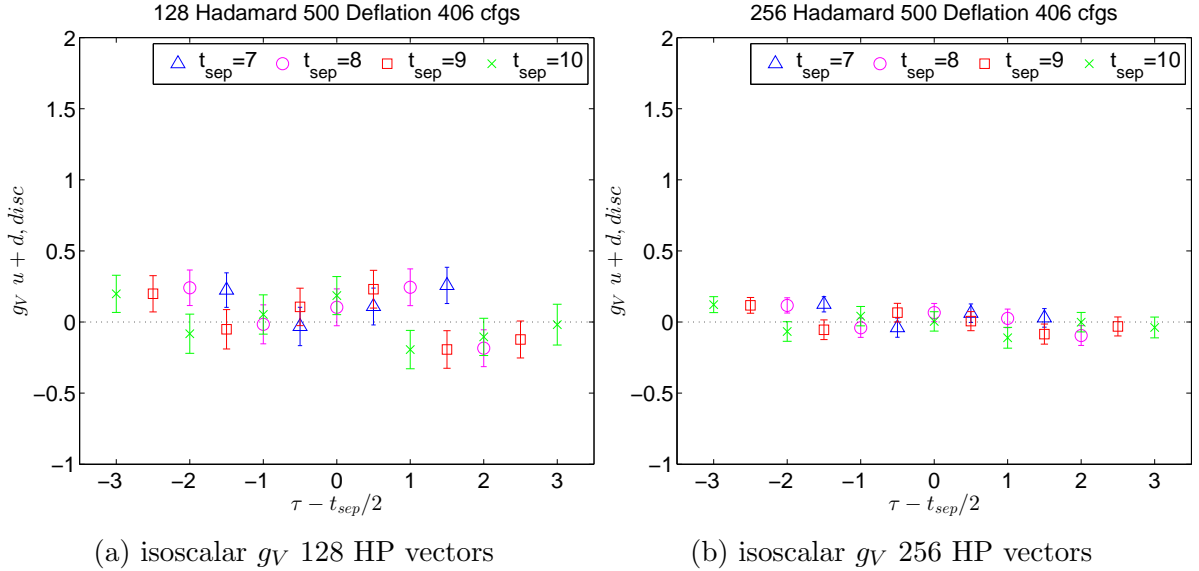


FIG. 6.17: A side by side comparison of the isoscalar disconnected vector charge at 128 and 256 HP vectors. Errors continue to be reduced and the expected answer is approached.

up to 256 HP vectors is still extremely beneficial. Based on the trend from 32-256 HP vectors, we suspect pushing to the third color closing point (512 HP vectors) would have reduced the errors even further. This is checked for the strange data, in which the runs were driven to 512 HP vectors.

We confirm with the strange data that running the trace estimator up to the 512 HP vectors still pays. This is of central importance to future studies of the vector current form factors. At 512 deflated HP vectors, the strange vector charge appears well pinned down. We conclude that the vector current does not fluctuate much over different field configurations, therefore a highly accurate trace is required to control the statistical errors.

In this section we have studied the axial, scalar, tensor, and vector charges as a function of the number of HP vectors. We determined that the axial and tensor charges require a reasonably accurate trace to saturate the statistical error. The scalar charge was found to be particularly insensitive to the error made by the trace estimator. From Figure 6.8, it is clear that even 32 deflated HP vectors may be an overkill. In contrast, the vector

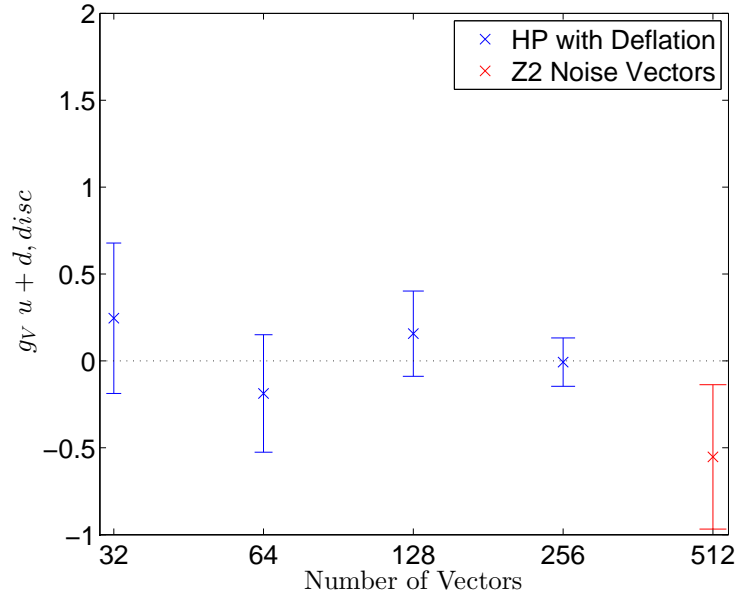


FIG. 6.18: The vector charge with deflated HP is compared to a pure noise trace estimator in the isoscalar sector. The points shown are with $t_{\text{sep}} = 10$, $\tau = 5$. Both methods are performed on the same 100 configurations.

charge necessitates an exceedingly accurate trace computation on each gauge configuration. Therefore, if loops are computed for the general purpose of hadron structure calculations, a large-scale and highly precise trace estimation cannot be avoided. This concludes our discussion of analyzing the dependence of final statistical errors on the noise of the trace estimator. In the next section, we perform fits with data from this and another ensemble to extract the ground state axial and scalar matrix elements.

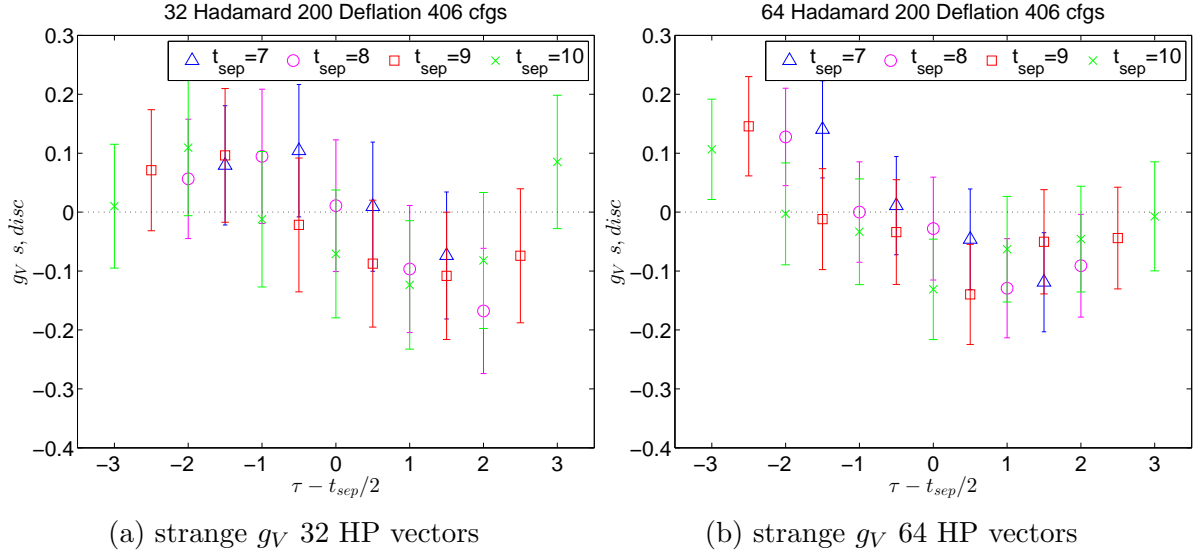


FIG. 6.19: A side by side comparison of the strange vector charge at 32 and 64 HP vectors. Errors reduce from 32 to 64 HP vectors.

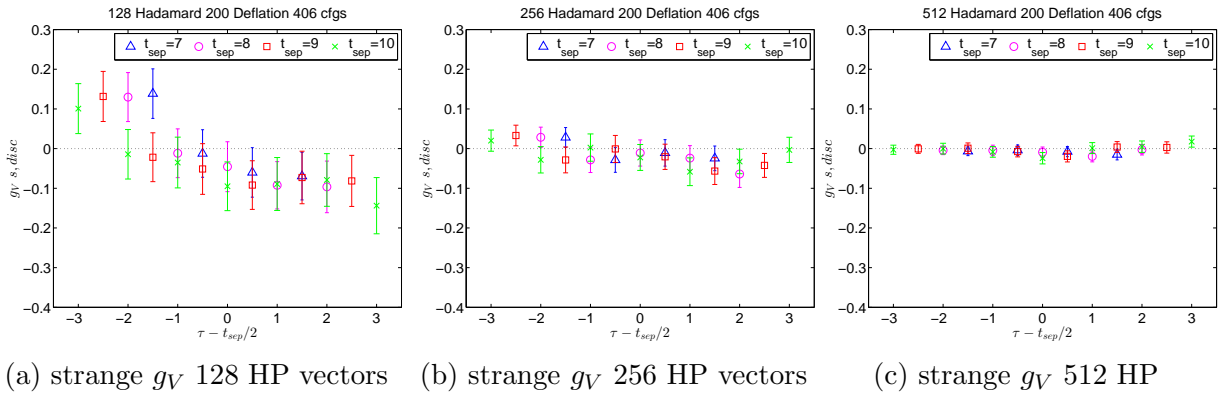


FIG. 6.20: A side by side comparison of the strange vector charge at 128, 256, and 512 HP vectors is given. We remark that the statistical error continues to go down, all the way through 512 HP vectors.

6.2 Disconnected nucleon matrix elements

Previously, we studied how the accuracy of the quark loop’s trace influences the statistical errors of different nucleon matrix elements. Here we use our most accurate runs and fit the ratio of three-point to two-point correlators to remove excited state contamination and isolate the ground state. Renormalized axial and scalar charges are presented at two different values of the lattice spacing.

We begin by introducing the relevant parameters from our runs on the two ensembles. Subsequently, fits to the data at multiple source-sink separations are presented, following the approach in Section 3.3. Results for zero-momentum transfer axial and scalar nucleon matrix elements are reported. Finally, we multiply our bare charges by the renormalization constants calculated from the NME collaboration [63].

6.2.1 Preliminaries

We compute quark loops on two ensembles, henceforth labeled as a127m285 and a094m280, corresponding to their respective values of lattice spacing and pion mass. Data in the varying HP investigation from the previous subsection were with loops computed on the a094m280 ensemble. Those same loops are used as part of this analysis, with the highest number of HP vectors taken. For the a127m285 ensemble, light loops were estimated with 512 HP and 500 SVD vectors. Strange loops on this ensemble were calculated with 512 HP vectors and no deflation vectors. Light loops were computed on 260 configurations of the a127m285 ensemble, whereas strange loops were computed on 528 configurations. We analyzed configurations separated by 10 updates on the a094m280 lattices. On the a127m285 lattices, configurations separated by 12 were taken for the light loops and a separation of 6 was used for the strange loops. Summaries of the two ensembles and loops are displayed in Tables 6.1 and 6.2.

TABLE 6.1: Ensemble Parameters

Ensemble	a in fm	m_π in MeV	β	c_{sw}	Volume	$m_\pi L$
a094m280	0.094	280	6.3	1.20536588	$32^3 \times 64$	4.11
a127m285	0.127	285	6.1	1.24930971	$32^3 \times 96$	5.85

TABLE 6.2: Quark Loop Parameters

Ensemble	Flavor	HP vectors	SVD Space	Number of cfgs
a094m280	light	256	500	406
a094m280	strange	512	200	406
a127m285	light	512	500	260
a127m285	strange	512	0	528

Just as for a094m280, we employ nucleon two-point correlation functions generated from the NME collaborations for the a127m285 ensemble. The spatial gauge links are hit with 20 iterations of stout smearing with a weight of $\rho = 0.08$ before the sources are constructed. Four equidistant timeslices are selected on the a127m286 lattice and 32 low-precision sources are inverted for each timeslice. The bias is then corrected for with one high-precision measurement. Smearing with a width of $\sigma = 5$ is used at the source and sink. Details of the two-point functions are given in Table 6.3. Note that when calculating the total number of measurements, a factor of 2 must be multiplied, resulting from forward and backward propagation. For example, the a094m280 ensemble has a total of $406 \times 96 \times 2 = 77952$ measurements.

TABLE 6.3: Nucleon Correlator Parameters

Ensemble	Smearing (Source, Sink)	Number of cfgs	Measurements per cfg
a094m280	(9,9)	406	96
a127m285	(5,5)	528	128

6.2.2 Axial charge

We fit the disconnected light and strange axial charge on both ensembles. We attempt fits to an exponential, the summation method, and the full two-state model, following the conventions in Section 3.3. Only fits with a reasonable reduced chi squared and sensible errors are shown. In the case of the exponential, only the middle insertion point is passed to the fitter. Fully covariant fits are taken when possible, however for ranges in which these fits are destabilized, we treat the data as uncorrelated instead. The data are jackknife resampled and the ratio of three to two-point correlators is fitted for each sample. For the axial matrix element, we are able to use the NME nucleon correlators polarized in the x , y , and z directions, giving a factor of 3 increase in statistics.

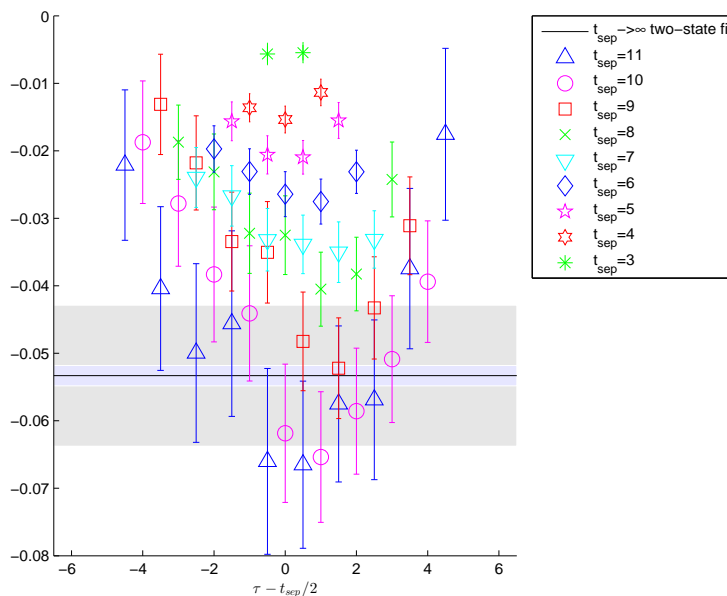


FIG. 6.21: A correlated two-state model fit of the disconnected light axial a094m280 data is taken. The fit is with a t_{sep} range of 3-11. The $\chi^2/d. o. f.$ is 1.34.

Figure 6.21 plots the ratio of the three-point to two-point correlator for the light axial matrix element on the a094m280 ensemble and Figures 6.22a and 6.22b on the a127m285 one. If the fit is correlated, the label “Covariant” is put before the fit name. We plot

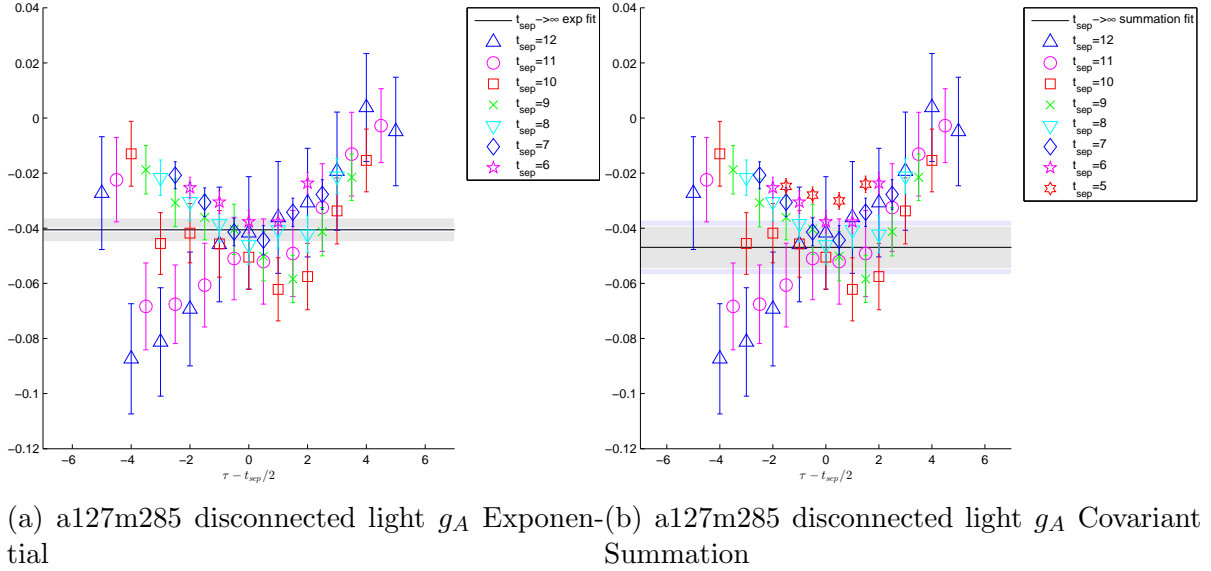


FIG. 6.22: The disconnected light axial data from the a127m285 ensemble are fitted with two different methods. The left plot is a fit to an exponential and the right is using the summation method. The exponential method has a source-sink separation fitting range of 6-12, while the summation fit employs a range of 5-12. The $\chi^2/\text{d.o.f.}$ are 1.37, and 1.43 respectively.

the fitting result of the ground state matrix element as a black line, corresponding to $t_{\text{sep}} \rightarrow \infty$. The gray band represents the statistical error, while the blue band gives the systematics. The systematical error is obtained by taking a variance of a particular fit with fits of neighboring t_{sep} ranges.

We are able to extract the disconnected light axial charge with small systematical uncertainties and a statistical error of roughly 20% for the a094m280 data and 10% for a127m285. The two fits performed on the a127m285 ensemble are in good agreement with each other. The magnitude of the central value of the axial charge decreases with coarser lattice spacing, however within the statistical uncertainty there is no difference. The strange axial data are also plotted in a similar fashion.

Figures 6.23a, 6.23b, and 6.23c show fitting results for exponential, summation, and two-state fits for the a094m280 strange axial data. There is nice agreement between the

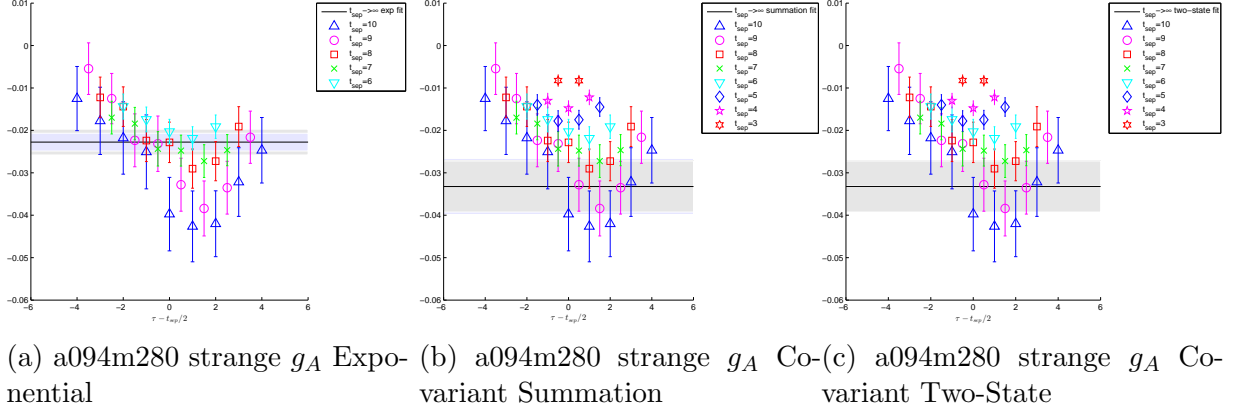


FIG. 6.23: The strange axial a094m280 data points are fitted with three different functional forms. The left plot is an exponential fit, the middle is with the summation method, and the right is to the two-state model. The exponential fit has a t_{sep} range of 6-10, the summation method and two-state model have ranges of 3-10. The $\chi^2/\text{d.o.f.}$ are 1.42, 1.25, and 1.05 respectively.

summation and two-state methods. The exponential fit yields a lower value of the strange axial charge and the errors are small enough to give a one σ discrepancy between the other fitting forms. We trust the two-state model result the most, due to its smaller reduced chi squared in spite of the fit utilizing the full covariance matrix. Furthermore, the systematical uncertainty of the two-state fit is two orders of magnitude smaller than its statistical error, which is why a light blue band is not visible in Figure 6.23c.

In Figure 6.24, a correlated two-state fit is taken for the strange axial charge with the a127m285 ensemble. Other fitting forms did not return serviceable results for this data. The axial charge is captured with approximately 20% error for this ensemble. The value matches the a094m280 result within error, demonstrating small lattice spacing errors with respect to the noise in our calculation.

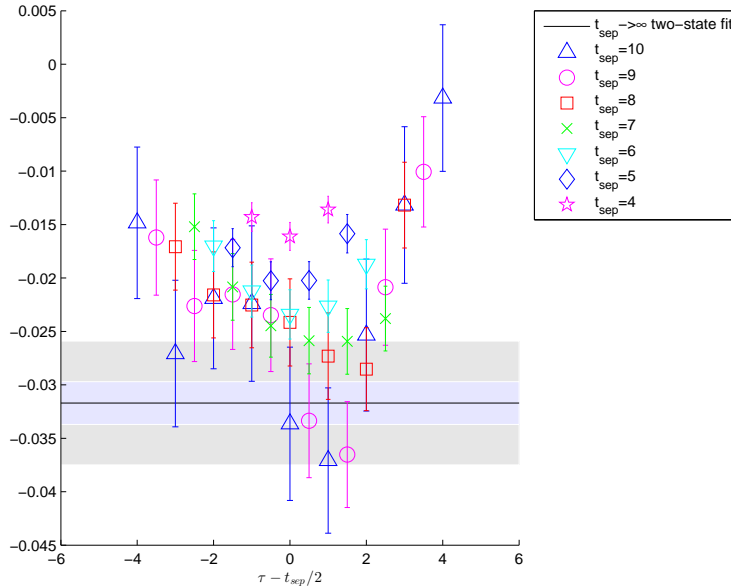
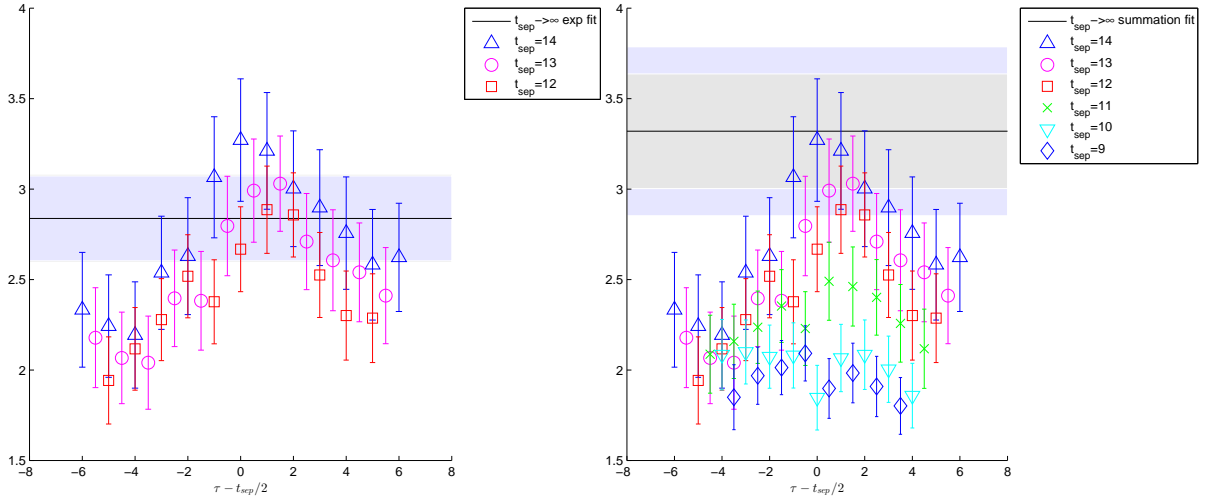


FIG. 6.24: The a127m285 strange axial data are fit to the two-state model from a t_{sep} range of 4-10. The fit is fully covariant. The $\chi^2/\text{d. o. f.}$ is 1.40.

6.2.3 Scalar charge

Here, the scalar matrix element is fitted with the same strategy as the previous subsection. Just as before, we plot the three-point correlator divided by the two-point one as a function of $t_{\text{sep}} - \tau$. This gives an “effective charge” that plateaus with large time separations, signaling little contamination from excited states. The fitted result of the scalar charge is plotted as an extrapolation of $t_{\text{sep}} \rightarrow \infty$, with statistical and systematical uncertainties denoted with gray and light blue bands correspondingly.

Figures 6.25a and 6.25b convey results of exponential and summation fits to the a094m280 light scalar ratio. There is reasonable agreement between them and we take the exponential result as our “golden value” due to lower statistical error and systematics. In Figure 6.26, the a127m285 fitting outcome is shown. The only successful fit is to the two-state model. The matrix element extracted from Figure 6.26 is much higher than in Figures 6.25a and 6.25b. The range of source-sink separations of the a127m285 two-state



(a) a094m280 disconnected light g_S Exponential (b) a094m280 disconnected light g_S Summation

FIG. 6.25: We plot the a094m280 disconnected light data with two fits, one to an exponential and one employing the summation method. The left plot has a t_{sep} range of 12-14, right one a range of 9-14. The $\chi^2/\text{d. o. f.}$ are 1.33 and 1.47 for the respective left and right figures.

fit was varied to check for a potentially misleading fit for a single range, however low systematics were found, as indicated in Figure 6.26. We also show data of the strange scalar charge from both ensembles.

Figures 6.27 and 6.28 show a similar story for the strange scalar charge as the light one. For each ensemble, only one fitting form proved effective; exponential for a094m280 and two-state for a127m285. Once again, the a127m285 matrix element is much above the a094m280 one. Both a127m285 fits were performed with the two-state model, it is possible that this model does not predict the ground state matrix element well for the scalar current. We were not able to produce two-state fits with reasonable reduced chi squares for the a094m280 ensemble, so a direct comparison between methods on the same data is not possible.

We point the reader to the next subsection, in which final answers for the renormalized charges are quoted. There, we will show that the ratio of light-to-strange scalar charge

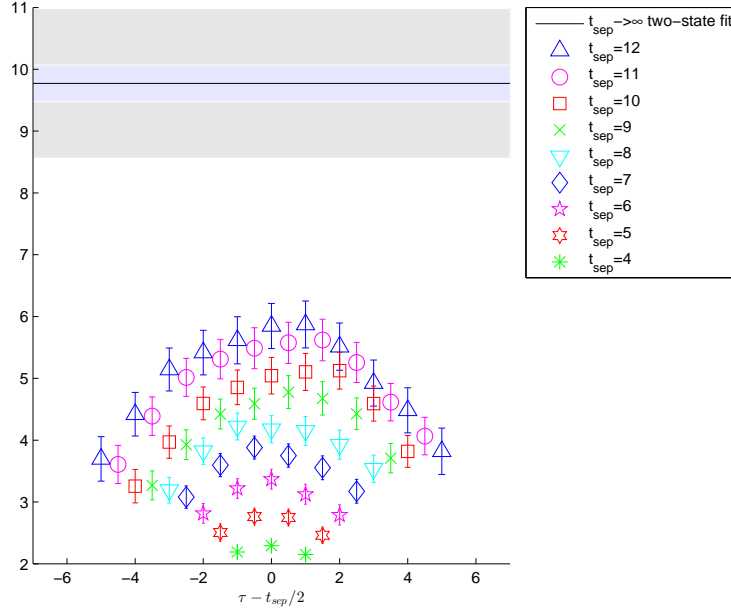


FIG. 6.26: The a127m285 light scalar points are fitted with the two-state model. The fit is fully covariant. The source-sink separations included are 4-12. The $\chi^2/\text{d. o. f.} = 1.38$.

is preserved well between the results of the two ensembles. Furthermore, the ratios of the a127m285 light to a094m280 light and a127m285 strange to a094m280 strange scalar charge are roughly equivalent.

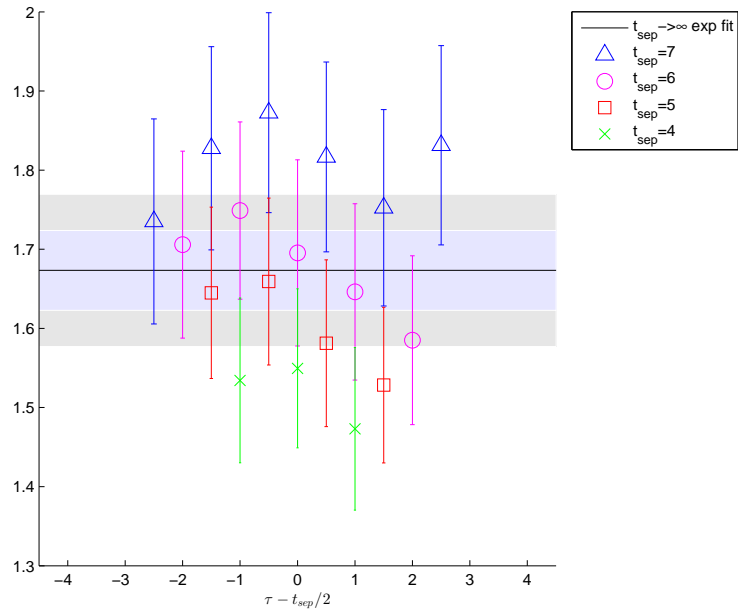


FIG. 6.27: The a094m280 strange scalar ratio is fitted to an exponential. The source-sink separations included are 4-7. The fit is uncorrelated and $\chi^2/\text{d.o.f.} = 0.94$.

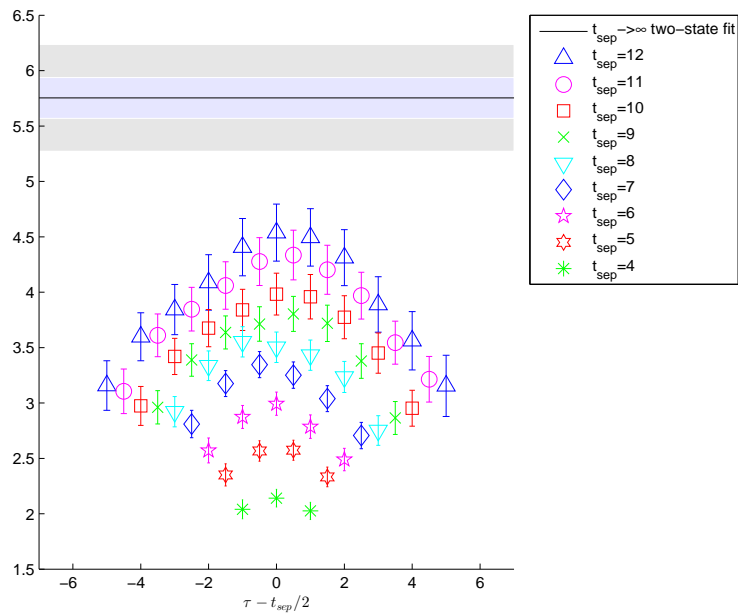


FIG. 6.28: The a127m285 light scalar points are fitted with the two-state model. The fit is fully covariant. The source-sink separations included are 4-12. The $\chi^2/\text{d.o.f.} = 1.38$.

6.2.4 Renormalized charges

Renormalization constants for both ensembles were computed by the NME collaboration in [63] using the RI-sMOM scheme matched to MS-bar at 2 GeV. We do not cover the RI-MOM scheme here and refer the interested reader to [72, 73]. It should be noted that these renormalization constants do not account for the mixing between light and strange axial currents, however a separate study of axial-vector form factors on the a127m285 ensemble have found this effect to be smaller than our statistical errors [74]. We choose fits with the least systematics for each charge and multiply by Z_A, Z_S to give the final results. Disconnected contributions to the light axial and scalar charges are given below.

$$\text{light } g_{A,a094m280}^{\text{two-state}} = -0.0473(93) \quad (6.4)$$

$$\text{light } g_{A,a127m285}^{\text{exp}} = -0.0357(30) \quad (6.5)$$

$$\text{light } g_{S,a094m280}^{\text{exp}} = 2.27(20) \quad (6.6)$$

$$\text{light } g_{S,a127m285}^{\text{two-state}} = 8.04(101) \quad (6.7)$$

The ensembles, labeled by lattice spacing and pion mass, are shown as subscripts and the fit selections are superscripts. The final error quoted is obtained through standard error propagation techniques on the error of the renormalization constant and corresponding charge. As mentioned in the previous text, the axial charges from the two ensembles are equal within error. The scalar charge shows a large difference, however. The matrix elements (6.5), (6.6), (6.7) have been computed with roughly 10% error; (6.4) shows 20% error. The strange axial and scalar renormalized charges are also computed on both ensembles.

$$\text{strange } g_{A,a094m280}^{\text{two-state}} = -0.0295(54) \quad (6.8)$$

$$\text{strange } g_{A,a127m285}^{\text{two-state}} = -0.0279(51) \quad (6.9)$$

$$\text{strange } g_{S,a094m280}^{\text{exp}} = 1.34(8) \quad (6.10)$$

$$\text{strange } g_{S,a127m285}^{\text{two-state}} = 4.74(41) \quad (6.11)$$

Similar to the light case, the strange axial charges between the two ensembles agree. Correspondingly, the strange scalar charge of the a127m285 ensemble is much larger than the a094m280 one. The strange axial charge values, (6.8) and (6.9) have both been obtained with a little under 20% error. The strange scalar charge, (6.11) has been obtained with 10% error and (6.10) with nearly 5% error.

If we compute the ratio of the light-to-strange scalar charge for the two different ensembles we get 1.694 and 1.696 respectively. Additionally, the ratios of the light-to-light and strange-to-strange scalar matrix elements under the two ensembles (with the a127m285 values as the numerator) are 3.541 and 3.537 respectively. These strange scalar fits were performed with different functional forms, additionally the nucleon two-point functions from the two ensembles involved different smearing. Although this is not a rigorous test, it serves as a consistency check in suggesting that the fits have captured the ground state matrix element well and some other effect is affecting light and strange flavors across the two ensembles. It is possible that this difference is due to unaccounted mixing between light, strange, and gluonic contributions. In [75, 76, 53] it was pointed out that the most straight forward calculation of the strange scalar charge from Wilson fermions produces larger results than with other lattice formulations due to explicit chiral symmetry breaking. This gives mixing between the light and strange scalar matrix elements for any finite lattice

spacing, even in the chiral limit. Our current method of renormalization does not account for this mixing, although a future study which does is forthcoming.

6.3 High-precision calculation of strange sigma terms in light nuclei

Now, we present fitting results to the strange sigma terms of the proton, proton-proton system, deuteron, and ${}^3\text{He}$ nucleus. The methodology for the quark loops computation matches the preceding calculations already discussed. We begin with a brief description of the lattice parameters of the calculation. Subsequently, we show fits for the zero-momentum transfer scalar current of the proton-proton, deuteron, and ${}^3\text{He}$, normalized by the scalar charge of the proton to avoid issues of renormalization. Finally, we organize results for the different fitting procedures, quoting both statistical and systematical uncertainty.

6.3.1 Lattice parameters and setup

Our analysis was carried out on a $32^3 \times 48$, $\beta = 6.1$ Clover improved Wilson ensemble with $c_{sw} = 1.24930970916466$. This ensemble is at the SU(3)-symmetric point, with a pion mass of $m_\pi \approx 805$ MeV. It should be noted that although we label the following calculations as the ratios of strange sigma terms, for this ensemble the results also represent the disconnected light sigma terms due to the degenerate masses of the up, down, and strange quarks at the SU(3)-symmetric point. The lattice spacing of this ensemble is $a \approx 0.12$ fm and $m_\pi L \approx 15.7$. One iteration of stout smearing with a weight of $\rho = .125$ was used in the fermion action. The measured configurations were separated by increments of 10 from the updates in gauge evolution. The hadron two-point functions were measured

over 416 forward and backward sources per configuration, on 508 configurations, resulting in a massive $508 \times 416 \times 2 = 422656$ number of measurements.

The large number of sources allowed us to overcome difficulties of large gauge noise in the multi-hadronic particles. Before generating the sources, the spatial links were hit with 10 iterations of stout smearing with $\rho = 0.08$. The hadron two-point functions were smeared at the source and sink with 80 iterations of Gaussian smearing with a width of $\sigma = 4.35$. The hadron three-point functions were computed using (6.3). Since the scalar matrix element has overlapping quantum numbers with the vacuum, the bubble diagram was subtracted. In the next subsections we show fits to the scalar charge of the proton, proton-proton, deuteron, and ${}^3\text{He}$ nucleus, labeled as “prot”, “pp”, “deut”, and “he3”.

6.3.2 Proton-proton

In order to extract the proton-proton matrix element, we do a jackknife resampling of the pp three and two-point correlators and fit each sample with the ratio method from Section 3.3. The same procedure is also done for the proton. Afterward, the ratio of the extracted pp and proton scalar matrix elements is computed for each sample. This final ratio is jackknife averaged, which gives the result we quote. Dividing the scalar charge of pp by the proton’s cancels the scalar renormalization constant. Moreover, calculating the ratio through resampling allows us to take advantage of correlations between the proton-proton and proton correlators, possibly canceling identical systematics appearing in both.

Since we are normalizing the multi-hadronic particle’s matrix element with the proton’s, a naive expectation value for the pp matrix element is 2. However, while the loop does in fact couple to twice as many quarks in the two-point correlator for pp in comparison to the proton, gluons are exchanged between these 6 quarks in a way that is not possible with two isolated protons. This creates a binding effect. Hence, we expect a value

slightly smaller than 2 due to this binding. The binding in the connected axial matrix elements of light nuclei has been studied in [77]. We use the same fitting methods that have been well established throughout this thesis.

TABLE 6.4: Proton-Proton/Proton Exponential Fit Result

Proton t_{sep}	PP t_{sep} 6 8	PP t_{sep} 7 8
4 8	1.952(38)	1.938(44)
5 8	1.918(34)	1.905(40)
7 8	1.895(32)	1.882(33)

TABLE 6.5: Proton-Proton/Proton Exponential Fit $\chi^2/d.o.f.$

Particle	t_{sep} 4 8	t_{sep} 5 8	t_{sep} 6 8	t_{sep} 7 8
Proton	1.17	0.47	N/A	0.35
PP	N/A	N/A	0.28	0.41

TABLE 6.6: Proton-Proton/Proton Summation Fit Result

Proton t_{sep}	PP t_{sep} 5 8	PP t_{sep} 6 8
6 8	1.937(72)	1.923(60)

Fitting results for the pp/proton ratio are given for three different fitting procedures and various source-sink separations in Tables 6.4, 6.6, and 6.8. Additionally, Tables 6.5, 6.7, and 6.9 display the reduced chi squared for each type of fit and t_{sep} range. If the fit is fully correlated, the “covariant” label prepends the fit name. We remark that only the two-state model results were reasonable under a covariant fit. Since the ratios are taken for each sample and then jackknife averaged, we allow the proton and pp fits to vary independently over different ranges of source-sink separation. Consequentially, we give the results as a matrix of ratios, where the rows vary the proton fitting range, while the columns vary the pp range. This matrix is not square, as some t_{sep} ranges may work for the proton, but not pp, and vice versa. It is expected that gauge noise in the proton will

TABLE 6.7: Proton-Proton/Proton Summation Fit $\chi^2/\text{d. o. f.}$

Particle	t_{sep} 5 8	t_{sep} 6 8
Proton	N/A	1.18
PP	0.70	1.12

TABLE 6.8: Proton-Proton/Proton Two-State Covariant Fit Result

Proton t_{sep}	PP t_{sep} 3 8	PP t_{sep} 4 8	PP t_{sep} 5 8	PP t_{sep} 6 8
3 8	1.879(37)	1.898(39)	1.904(39)	1.877(52)
4 8	1.856(39)	1.875(36)	1.881(37)	1.854(49)
5 8	1.858(41)	1.877(39)	1.883(35)	1.856(47)
6 8	1.831(66)	1.849(64)	1.856(63)	1.829(56)
7 8	1.775(99)	1.793(98)	1.799(97)	1.773(92)

be significantly less than in pp. Fitting ranges that are successful on one particle but not the other are labeled as N/A in Tables 6.5, 6.7, and 6.9 for the respective case in which they fail.

Inspecting Tables 6.4, 6.6, and 6.8, we find good agreement with our hypothesis of a value slightly lower than 2 due to binding. For each fitting type, we select the t_{sep} ranges in pp and the proton that give a matrix element closest to all the others for that particular form. This number is taken as the “representative” value of that fit and a systematic is assigned based on the variance of that value with all the other ranges. Since only source-sink separation ranges that produced reasonable results in terms of the reduced chi squared and statistical error are shown, we find this as the most honest way to allot a systematical uncertainty to our results.

The t_{sep} ranges chosen for the proton are 5-8, 6-8, and 6-8 for the exponential, summation, and two-state fits respectively. The separations for pp are 6-8, 6-8, and 4-8, correspondingly. The final ratios, based on these separations, are quoted in Section 6.3.5, with systematical and statistical uncertainties. The results are also plotted. Since there is a matrix of ratios for each fitting form, we plot only the t_{sep} dependence of pp, keeping

TABLE 6.9: Proton-Proton/Proton Two-State Covariant Fit $\chi^2/\text{d. o. f.}$

Particle	t_{sep} 3 8	t_{sep} 4 8	t_{sep} 5 8	t_{sep} 6 8	t_{sep} 7 8
Proton	1.22	1.22	1.01	0.84	0.85
PP	1.22	1.18	1.13	1.11	N/A

the proton t_{sep} fixed to the values given above.

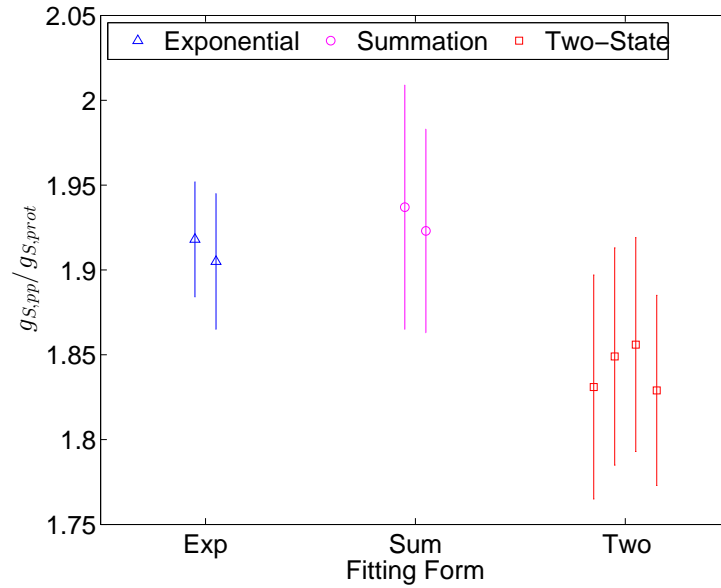


FIG. 6.29: The ratio of g_S for the proton-proton divided by the proton is plotted. The proton t_{sep} is fixed at 5-8 for the exponential data and 6-8 for the summation method and two-state model results. The t_{sep} ranges of pp follow from Tables 6.4, 6.6, and 6.8.

In Figure 6.29, within statistical uncertainties, we see roughly equivalent values of the scalar charge ratio between the three methods. The different data points for each fitting type follow from the columns of their corresponding tables. The two exponential points (blue) are from ranges 6-8 and 7-8, the two summation points (magenta) are from 5-8 and 6-8, and the two-state model points (red) cover 3-8, 4-8, 5-8, and 6-8. The errors for all points plotted are under 10%.

6.3.3 Deuteron

We repeat the same procedure from the former subsection for the deuteron. The ratio of the deuteron three-point to two-point correlators is fitted on each jackknife sample. The proton scalar charge is also extracted similarly. The ratio of the two scalar charges is taken for each sample and then averaged. Once again, the expectation for this ratio is a little less than 2 for the deuteron.

TABLE 6.10: Deuteron/Proton Exponential Fit Result

Proton t_{sep}	Deut t_{sep} 4 8	Deut t_{sep} 6 8	Deut t_{sep} 7 8
3 8	1.950(25)	2.013(45)	2.007(53)
4 8	1.881(22)	1.941(39)	1.935(47)
5 8	1.848(23)	1.907(34)	1.902(42)
7 8	1.826(32)	1.884(32)	1.879(35)

TABLE 6.11: Deuteron/Proton Exponential Fit $\chi^2/d. o. f.$

Particle	t_{sep} 4 8	t_{sep} 5 8	t_{sep} 6 8	t_{sep} 7 8
Proton	1.17	0.47	N/A	0.35
Deut	0.84	N/A	0.30	0.67

TABLE 6.12: Deuteron/Proton Summation Fit Result

Proton t_{sep}	Deut t_{sep} 4 8	Deut t_{sep} 5 8	Deut t_{sep} 6 8
6 8	1.956(65)	1.929(60)	1.914(61)

Tables 6.10, 6.12, 6.14 summarize results of the fits to the deuteron. The respective reduced chi squares are given in Tables 6.11, 6.13, 6.15. Only the two-state model fits take into account off-diagonal elements of the covariance matrix. We observe relatively similar values to that of the pp ratio. As expected, the matrix elements are slightly below 2.

There is a greater range of source-sink separations with converging fits for the deuteron exponential fit compared to the pp one (see Tables 6.4 and 6.10). However, the deuteron

TABLE 6.13: Deuteron/Proton Summation Fit $\chi^2/\text{d. o. f.}$

Particle	t_{sep} 4 8	t_{sep} 5 8	t_{sep} 6 8
Proton	N/A	N/A	1.18
Deut	0.70	0.92	1.51

TABLE 6.14: Deuteron/Proton Two-State Covariant Fit Result

Proton t_{sep}	Deut t_{sep} 3 8	Deut t_{sep} 4 8	Deut t_{sep} 5 8
3 8	1.884(42)	1.913(49)	1.916(49)
4 8	1.861(44)	1.890(44)	1.892(44)
5 8	1.863(46)	1.892(47)	1.894(42)
6 8	1.836(66)	1.864(63)	1.866(59)
7 8	1.780(99)	1.807(97)	1.810(94)

exponential fits fluctuate more, giving larger systematics. On the other hand, there are more acceptable two-state model fits for the proton-proton as opposed to the deuteron, seen in Tables 6.8 and 6.14. The two-state covariant fit appears to have smaller systematics than the statistical errors.

We choose a single representative point of the fits, as described in Section 6.3.2 and arrive at t_{sep} ranges of 5-8, 6-8, 6-8 for the proton, corresponding to exponential, summation, and two-state results respectively. The deuteron source-sink ranges are 6-8, 5-8, and 4-8 correspondingly. We remind the reader that full results, including the systematics, will be shown in the last subsection. The ratio of deuteron to proton scalar charges using the static proton t_{sep} ranges from above is plotted.

Figure 6.30 displays agreement of the scalar charge ratio between the three methods within errors. The exponential points (blue) are from ranges 4-8, 6-8 and 7-8, the two summation points (magenta) are from 4-8, 5-8 and 6-8, and the two-state model points (red) go from 3-8, 4-8, and 5-8. In both Figures 6.29 and 6.30, the summation form produces somewhat higher central values than the exponential and two-state methods.

TABLE 6.15: Deuteron/Proton Two-State Covariant Fit $\chi^2/\text{d. o. f.}$

Particle	t_{sep} 3 8	t_{sep} 4 8	t_{sep} 5 8	t_{sep} 6 8	t_{sep} 7 8
Proton	1.22	1.22	1.01	0.84	0.85
Deut	1.11	1.05	1.04	N/A	N/A

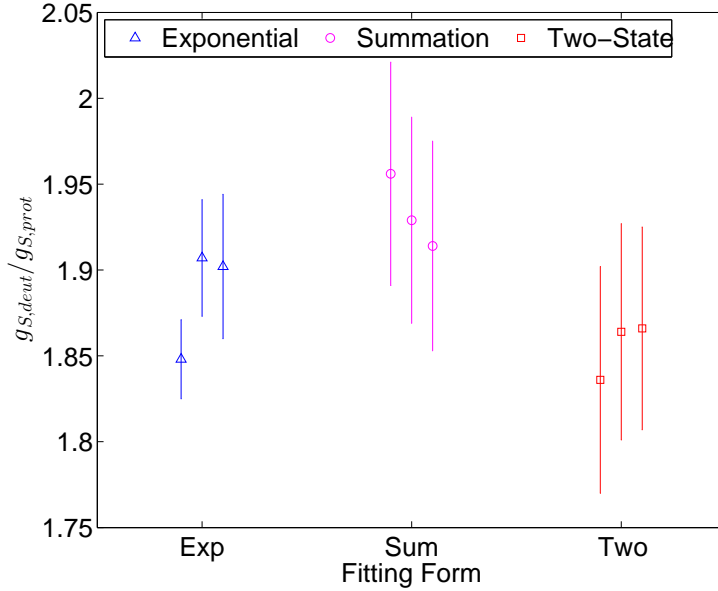


FIG. 6.30: The ratio of g_S for the deuteron to proton ratio is given. The proton t_{sep} is fixed at 5-8 for the exponential data and 6-8 for the summation method and two-state model results. The t_{sep} ranges of the deuteron are the columns of Tables 6.10, 6.12, and 6.14.

6.3.4 ${}^3\text{He}$

Finally, the zero-momentum transfer ${}^3\text{He}$ scalar matrix element is analyzed. Similar to the other particles, we divide by the proton scalar charge. Since a ${}^3\text{He}$ particle is made up of 3 hadrons, the naive expectation for its ratio with the proton is 3. We hypothesize a value a little less than 3 because of binding.

In Tables 6.16, 6.18, 6.20 the extracted scalar charge ratio is given and in Tables 6.17, 6.19, 6.21 the corresponding reduced chi squares are shown. Just as the other cases, the two-state model data are the only results from correlated fits. The ${}^3\text{He}$ scalar charge divided by proton charge is a bit less than 3, confirming our expectations.

TABLE 6.16: ${}^3\text{He}/\text{Proton}$ Exponential Fit Result

Proton t_{sep}	He3 t_{sep} 4 8	He3 t_{sep} 6 8	He3 t_{sep} 7 8
4 8	2.640(246)	2.727(115)	2.645(133)
5 8	2.594(242)	2.680(110)	2.600(128)
7 8	2.563(237)	2.648(108)	2.569(123)

TABLE 6.17: ${}^3\text{He}/\text{Proton}$ Exponential Fit $\chi^2/\text{d. o. f.}$

Particle	t_{sep} 4 8	t_{sep} 5 8	t_{sep} 6 8	t_{sep} 7 8
Proton	1.17	0.47	N/A	0.35
He3	0.59	N/A	0.65	0.67

The gauge noise of the ${}^3\text{He}$ matrix elements is higher than that of the other two particles. Additionally, the binding appears stronger than in the proton-proton or deuteron. The greatest number of t_{sep} ranges for the two-state covaraint fit are achieved with ${}^3\text{He}$. However, the numbers drift downward as the minimum proton t_{sep} is increased. Unlike the previous two particles, the two-state model suffers from larger systematical uncertainties than the other two methods. This can easily be seen visually, by comparing Table 6.20 to Tables 6.16 and 6.18; we will give concrete values for the systematics in the forthcoming subsection.

Just as for pp and the deuteron, a representative point is selected for each functional fitting form. The proton t_{sep} ranges are 4-8, 6-8, 6-8, while the ${}^3\text{He}$ ranges are 4-8, 6-8, 3-8, corresponding to the exponential, summation, and two-state methods. We plot the ${}^3\text{He}$ scalar matrix element using these values for the proton t_{sep} . The different points in the same color represent the varying ${}^3\text{He}$ source-sink separations

In Figure 6.31, we find equivalent answers within errors for all three fitting procedures. The central values of the exponential and summation methods are close, the two-state fit yields a lower ratio. Furthermore, the 7-8 point of the two-state fit (the last red point) drops compared to the others, leading to a higher systematical uncertainty. ${}^3\text{He}$ is the only

TABLE 6.18: ${}^3\text{He}/\text{Proton}$ Summation Fit Result

Proton t_{sep}	He3 t_{sep} 5 8	He3 t_{sep} 6 8
6 8	2.718(177)	2.714(224)

TABLE 6.19: ${}^3\text{He}/\text{Proton}$ Summation Fit $\chi^2/\text{d.o.f.}$

Particle	t_{sep} 5 8	t_{sep} 6 8
Proton	N/A	1.18
He3	0.99	1.46

particle for which a stable fit was found for the 7-8 range. It is possible that although the reduced chi squared for this range is reasonable, there is not enough data to fully capture all the fitting parameters well, leading to an underestimate of the matrix element. Since the chi squared is sensible however, we leave this point in when computing a “systematical variance”. Final results for all the charges with systematical and statistical uncertainties are given in the next subsection.

TABLE 6.20: ${}^3\text{He}/\text{Proton}$ Two-State Covariant Fit Result

Proton t_{sep}	He3 t_{sep} 3 8	He3 t_{sep} 4 8	He3 t_{sep} 5 8	He3 t_{sep} 6 8	He3 t_{sep} 7 8
3 8	2.568(98)	2.609(100)	2.638(95)	2.604(137)	2.384(82)
4 8	2.536(99)	2.577(98)	2.606(93)	2.572(132)	2.355(85)
5 8	2.539(99)	2.580(99)	2.609(92)	2.575(130)	2.357(86)
6 8	2.502(121)	2.542(119)	2.571(115)	2.537(132)	2.323(116)
7 8	2.426(158)	2.465(158)	2.493(155)	2.460(166)	2.252(152)

 TABLE 6.21: ${}^3\text{He}/\text{Proton}$ Two-State Covariant Fit $\chi^2/\text{d. o. f.}$

Particle	t_{sep} 3 8	t_{sep} 4 8	t_{sep} 5 8	t_{sep} 6 8	t_{sep} 7 8
Proton	1.22	1.22	1.01	0.84	0.85
He3	1.25	1.17	1.15	1.20	1.19

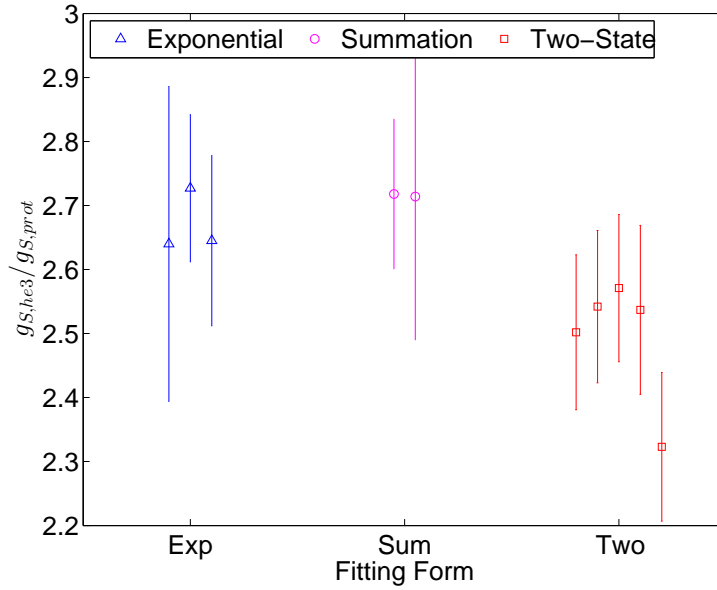


FIG. 6.31: The scalar charge of ${}^3\text{He}$ divided by the scalar charge of the proton is plotted. The proton t_{sep} is 4-8 for the exponential data and 6-8 for the summation method and two-state model. The t_{sep} ranges of the ${}^3\text{He}$ charge are the columns of Tables 6.16, 6.18, and 6.20.

6.3.5 Final results

We quote final estimates and errors for the ratios of strange g_S by selecting fit ranges which are closest to all other values for a particular form. For the uncorrelated exponential fits, we choose the t_{sep} ranges of 6-8 for proton-proton, 6-8 for the deuteron, and 4-8 for ^3He . The corresponding ranges for the proton portion of the fits are 5-8, 5-8, and 4-8 respectively.

$$g_{S,pp}^{\text{exp}}/g_{S,\text{prot}}^{\text{exp}} = 1.918(30)(38) \quad (6.12)$$

$$g_{S,\text{deut}}^{\text{exp}}/g_{S,\text{prot}}^{\text{exp}} = 1.907(61)(34) \quad (6.13)$$

$$g_{S,\text{he3}}^{\text{exp}}/g_{S,\text{prot}}^{\text{exp}} = 2.640(58)(246) \quad (6.14)$$

Above, the first error is due to systematic uncertainty and was computed by taking a variance of the selected results with all other values shown in Tables 6.4, 6.10, and 6.16. The square root of this variance is the error shown. The second error is the statistical uncertainty. The systematics of the exponential fit is small for the proton-proton and ^3He . As Figure 6.30 confirms, the systematics of the deuteron exponential fit were large. Final estimates of error from the uncorrelated summation fits were also computed.

$$g_{S,pp}^{\text{summation}}/g_{S,\text{prot}}^{\text{summation}} = 1.923(14)(60) \quad (6.15)$$

$$g_{S,\text{deut}}^{\text{summation}}/g_{S,\text{prot}}^{\text{summation}} = 1.929(31)(60) \quad (6.16)$$

$$g_{S,\text{he3}}^{\text{summation}}/g_{S,\text{prot}}^{\text{summation}} = 2.714(004)(224) \quad (6.17)$$

For the summation fits, t_{sep} ranges of 6-8, 5-8, and 6-8 were taken for proton-proton,

deuteron, and ${}^3\text{He}$ respectively. The matching proton fit ranges were all 6-8. Although the statistical errors are slightly higher than in the exponential case, systematics of the summation fits were incredibly small for all three particles.

Finally, systematics and statistical errors for the correlated fits with the full two-state model are given. The proton t_{sep} range was 6-8 for all 3 cases. The fit ranges for the multi-hadronic particles were 4-8 for the proton-proton, 4-8 for the deuteron, and 3-8 for ${}^3\text{He}$. Since the two-state model was the only fitting method that take full advantage of the correlations in our data, we expect that it gives the most honest assessment of the errors in our calculation. Systematics for the proton-proton and deuteron were under the statistical error. Furthermore, for the two-hadron particles, the errors are completely compatible with the summation method results. The ${}^3\text{He}$ matrix element had a larger systematic, as seen in Table 6.20. Nevertheless, its value is consistent with the other fitting methods within error.

$$g_{S,pp}^{\text{two-state}} / g_{S,prot}^{\text{two-state}} = 1.849(40)(64) \quad (6.18)$$

$$g_{S,deut}^{\text{two-state}} / g_{S,prot}^{\text{two-state}} = 1.864(42)(63) \quad (6.19)$$

$$g_{S,he3}^{\text{two-state}} / g_{S,prot}^{\text{two-state}} = 2.568(106)(98) \quad (6.20)$$

Although systematics varied, all three fitting forms produced compatible results within the statistical uncertainty and systematical error we have assigned. We are able to compute the proton-proton and deuteron ratios between 3-4%. The ${}^3\text{He}$ scalar charge was obtained with an uncertainty of under 10% in all the fits. We are not aware of any other calculations of disconnected diagrams in light nuclei and believe this computation is the first. Furthermore, due to our massive number of measurements and novel algorithms, this calculation is extremely precise.

CHAPTER 7

Summary

In this work, we have introduced improved methods to compute the trace of a matrix function and applied them to disconnected diagrams in Lattice QCD. We have studied the effects of singular value deflation on trace estimation theoretically and experimentally on a variety of matrices, both from QCD and from other fields. The synergistic combination of hierarchical probing and deflation was used to compute light quark disconnected nucleon observables on ensembles with two different lattice spacings. Additionally, these algorithms were applied to the first calculation of strange sigma terms in light nuclei.

In our analysis of the effects of deflation, we introduced a model from [56] that requires only the spectrum of singular values as input to capture the qualitative features of variance reduction under complex Rademacher noise. The same examination was performed with Gaussian noise, although no model was required and an exact study could be conducted. We have demonstrated the robustness of our model and the rare property of non-Hermitian matrices performing better under deflation than Hermitian ones. Additionally, a comparison with deflated Gaussian noise and \mathbb{Z}_4 noise has been investigated. For the matrices tested, we have found deflated \mathbb{Z}_4 to always perform better, be it sometimes

marginally or substantially.

Additionally, the synergistic role between hierarchical probing and deflation was explored on matrices from Lattice QCD production runs. Over a factor of 60 variance reduction was demonstrated when comparing to the basic Monte Carlo trace method. Varying the size of the singular value space revealed an “optimal” number of deflation vectors to use for a given color closing point. By totaling wallclock of our eigensolver and trace estimator, we show little overhead for generating the deflation subspace.

We have applied deflated hierarchical probing to compute nucleon/nuclear matrix elements using clover-improved Wilson fermions at pion masses ranging from ≈ 300 MeV to 800 MeV. The accuracy of the trace per configuration was studied for nucleon axial, scalar, tensor, and vector current matrix elements. We have found an extremely accurate trace is required to minimize noise for the vector and tensor charges. The axial matrix element necessitates a moderate accuracy of the trace and the scalar matrix element saturates with a very small number of probing vectors.

An analysis of controlling excited state contamination and isolating the ground state of the disconnected nucleon axial and scalar matrix elements followed. We obtained results of $-0.0295(54)$ and $-0.0279(51)$ for the strange axial charge on two ensembles of varying lattice spacing, with the caveat that our renormalization procedure did not account for mixing, although this effect is expected to be below our statistical errors. The current value of the strange axial charge from polarized deep inelastic scattering data is $-0.03(03)$; our results show agreement and higher precision.

Finally, we computed the ratio of strange sigma terms in light nuclei to that of the proton’s. These ratios are $1.849(40)(64)$, $1.864(42)(63)$, and $2.568(106)(98)$ for the proton-proton, deuteron, and ${}^3\text{He}$ nucleus respectively, where the first error is systematical and the second is statistical. The results are quoted based on the two-state model, however other fitting methods yielded compatible values within error. We note the uncertainty of

our calculation is small enough to be able to observe a significant binding effect of the strange sigma terms in light nuclei.

APPENDIX A

Euclidean Gamma matrices

The Minkowski gamma matrices obey the Clifford algebra

$$\{\gamma_\mu, \gamma_\nu\} = 2g_{\mu\nu}, \quad (\text{A.1})$$

where the metric tensor is

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Therefore, a Euclidean version of the γ matrices may be defined with the following rotation:

$$\begin{aligned}
\gamma_1^E &= -i\gamma_1 \\
\gamma_2^E &= -i\gamma_2 \\
\gamma_3^E &= -i\gamma_3 \\
\gamma_4^E &= \gamma_0.
\end{aligned}$$

This yields new anti-commutation relations, in which the metric is the identity.

$$\{\gamma_\mu^E, \gamma_\nu^E\} = 2\delta_{\mu\nu} \quad (\text{A.2})$$

These Euclidean γ_μ^E matrices also allow for a γ_5^E to be defined based on the product

$$\gamma_5^E = \gamma_1^E \gamma_2^E \gamma_3^E \gamma_4^E. \quad (\text{A.3})$$

It can easily be checked that γ_5^E anti-commutes with all the other γ_μ^E matrices, and that $(\gamma_5^E)^2 = \mathbb{1}$. In the chiral representation, γ_5^E is diagonal, and the γ_μ^E matrices take the form of

$$\begin{aligned}
\gamma_1^E &= \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \gamma_2^E = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \gamma_3^E = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \\ i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix}, \\
\gamma_4^E &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \gamma_5^E = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.
\end{aligned}$$

In this explicit form it can also be readily verified that

$$\gamma_\mu^E = (\gamma_\mu^E)^\dagger = (\gamma_\mu^E)^{-1}. \quad (\text{A.4})$$

More details of γ matrices can be found in [78].

APPENDIX B

Grassmann numbers

It is well known from Fermi statistics that the product of two fermion fields must be antisymmetric with respect to their interchange.

$$\psi(m)_{\alpha,a}\psi(n)_{\beta,b} = -\psi(n)_{\beta,b}\psi(m)_{\alpha,a} \quad (\text{B.1})$$

Equation B.1 is true for fermions of different flavors and also of ψ and $\bar{\psi}$. Therefore, the fermionic degrees of freedom of any quantum field theory must be described by anti-commuting numbers, or Grassmann numbers. Two Grassmann numbers satisfy the general property

$$\theta_1\theta_2 = -\theta_2\theta_1. \quad (\text{B.2})$$

This indicates that the square of any Grassmann number is zero, $\theta_1^2 = \theta_2^2 = 0$. Consequently, the space of possible functions is not very rich. Only linear and zeroth order polynomials with respect to any Grassmann variable are permitted. The most general polynomial that may be written for two Grassmann numbers is

$$P = a + b_1\theta_1 + b_2\theta_2 + c\theta_1\theta_2, \quad (\text{B.3})$$

where a, b, c, d are complex coefficients. Additionally, we introduce a partial derivative with respect to θ_1 ,

$$\frac{\partial}{\partial\theta_1}P = b_1 + c\theta_2. \quad (\text{B.4})$$

However, since the last term in (B.3) can be rewritten as $-c\theta_2\theta_1$, we conclude that the derivative must also act like a Grassmann number in order to produce consistent results.

$$\frac{\partial}{\partial\theta_1}\theta_2 = -\theta_2\frac{\partial}{\partial\theta_1} \quad (\text{B.5})$$

$$\frac{\partial}{\partial\theta_1}\frac{\partial}{\partial\theta_2} = -\frac{\partial}{\partial\theta_2}\frac{\partial}{\partial\theta_1} \quad (\text{B.6})$$

$$\frac{\partial^2 P}{\partial\theta_1^2} = \frac{\partial^2 P}{\partial\theta_2^2} = 0 \quad (\text{B.7})$$

Next we consider integration under Grassmann variables. For simplicity, we focus temporarily on a function of only one Grassmann number.

$$f(\theta_1) = a + b\theta_1 \quad (\text{B.8})$$

If $f(\theta_1)$ is integrated with respect to θ_1 over all space, the integral must be invariant under shifts by a constant, $\theta_1 \rightarrow \theta_1 + \theta_2$.

$$\int d\theta_1(a + b\theta_1) = \int d\theta_1(a + b\theta_2 + b\theta_1) \quad (\text{B.9})$$

Since θ_2 is arbitrary, this gives

$$\int d\theta_1 = 0 \tag{B.10}$$

$$\int d\theta_1 \theta_1 = 1, \tag{B.11}$$

where in the second line we have enforced a normalization to give an answer proportional to b for $\int f(\theta_1)$. Notice that differentiation and integration with respect to a Grassmann number are identical. It is trivial to extend the definition of the integral to several variables.

$$\int d^N \theta \theta_1 \theta_2 \dots \theta_N = 1 \tag{B.12}$$

$$\int d^N \theta = \int d\theta_N d\theta_{N-1} \dots d\theta_1 \tag{B.13}$$

Using these definitions we study a specific, but useful two-dimensional integral, a Gaussian. This is evaluated easily by expanding the exponential in a Taylor series and recalling that only up to linear Grassmann polynomials are non-vanishing.

$$\int d\bar{\theta} d\theta e^{-\bar{\theta} a \theta} = \int d\bar{\theta} d\theta (1 - \bar{\theta} a \theta) = a \tag{B.14}$$

Let us now consider a Gaussian integral with two pairs of θ and $\bar{\theta}$.

$$\int d\theta_1 d\bar{\theta}_1 d\theta_2 d\bar{\theta}_2 e^{-\bar{\theta}_1 a_1 \theta_1} e^{-\bar{\theta}_2 a_2 \theta_2} = a_1 a_2 = e^{\log a_1 + \log a_2} \tag{B.15}$$

The result $a_1 a_2$ may be viewed as the exponential sum of logs, or the determinant of a diagonal matrix whose non-zero elements are a_1 and a_2 . This gives a generalization to Gaussian integration over many Grassmannian degrees of freedom.

$$\int \prod_k^N d\theta_k d\bar{\theta}_k e^{-\bar{\theta}_i A_{ij} \theta_j} = e^{\text{Tr} \log A} = \det A \quad (\text{B.16})$$

Finally, this result may be applied to (3.10).

$$Z_F[U] = \int \mathcal{D}[\psi, \bar{\psi}] e^{-S_F[\psi, \bar{\psi}, U]} = \int \mathcal{D}[\psi, \bar{\psi}] e^{-\bar{\psi} D[U] \psi} = \det D[U] \quad (\text{B.17})$$

Above, we have suppressed all indices in favor of matrix/vector notation. Hence $Z_F[U]$ is known as the fermion determinant.

APPENDIX C

A closer look at conjugate gradient

The j th search direction in the conjugate algorithm is determined by

$$p_{j+1} = r_j - \sum_{i=0}^j \beta_{i,j} p_i \tag{C.1}$$

$$\beta_{i,j} = \frac{p_i^\dagger A r_j}{p_i^\dagger A p_i}. \tag{C.2}$$

However, the final form in Algorithm 3 reduces this sum to only the last term. Furthermore, the expression for β_j looks different from above and $\beta_j p_j$ is added instead of subtracted from r_j to compute p_{j+1} . To resolve these discrepancies, we begin by noting that the current residual r_j is

$$r_j = Ax - Ax_j = \sum_{i=j+1}^N \alpha_i A p_i. \tag{C.3}$$

Multiplying both sides by the dual of an arbitrary search direction p_k , gives

$$p_k^\dagger r_j = \sum_{i=j+1}^N \alpha_i p_k^\dagger A p_i. \quad (\text{C.4})$$

For $k \leq j$, all terms in the right-hand sum will be zero. This implies that $p_k^\dagger r_j = 0$ under the same condition, put simply the current residual is orthogonal to all previous search directions. Furthermore, since previous residuals are made up of a linear combination of previous search directions, this also suggests that the current residual is orthogonal to all previous ones. We may exploit this orthogonality to derive the expression for α_j used in Algorithm 3.

$$\begin{aligned} r_{j+1}^\dagger r_j &= 0 \\ (b - Ax_{j+1})^\dagger r_j &= 0 \\ [b - A(x_j + \alpha_{j+1} p_{j+1})]^\dagger r_j &= 0 \\ r_j^\dagger r_j - \alpha_{j+1} (A p_{j+1})^\dagger r_j &= 0 \\ \alpha_{j+1} &= \frac{r_j^\dagger r_j}{p_{j+1}^\dagger A r_j} \end{aligned} \quad (\text{C.5})$$

We may substitute (C.1) for r_j and remind the reader that all p_i in the sum are conjugate to p_{j+1} .

$$\alpha_{j+1} = \frac{r_j^\dagger r_j}{p_{j+1}^\dagger A p_{j+1}} \quad (\text{C.6})$$

Now all the ingredient are in place to demonstrate why all $\beta_{i,j}$ vanish for $i < j$. We start by relating the next residual to the previous one.

$$\begin{aligned}
r_{j+1} &= r_j - \alpha_{j+1} A p_{j+1} \\
r_i^\dagger r_{j+1} &= r_i^\dagger r_j - \alpha_{j+1} r_i^\dagger A p_{j+1}
\end{aligned}$$

Subsequently, the definitions for β and α are plugged into the right-hand side.

$$\begin{aligned}
r_i^\dagger r_{j+1} &= r_i^\dagger r_j - \alpha_{j+1} \beta_{j+1,i} p_{j+1}^\dagger A p_{j+1} \\
r_i^\dagger r_{j+1} &= r_i^\dagger r_j - \beta_{j+1,i} r_j^\dagger r_j \\
\beta_{j+1,i} &= \frac{r_i^\dagger r_j - r_i^\dagger r_{j+1}}{r_j^\dagger r_j} \\
\beta_{i,j} &= \frac{r_j^\dagger r_{i-1} - r_j^\dagger r_i}{r_{i-1}^\dagger r_{i-1}} \tag{C.7}
\end{aligned}$$

In the last step we have relabeled indices to obtain a simple expression for $\beta_{i,j}$. It is clear now that for $i < j$, the numerator of (C.7) is always 0 due to the orthogonality of the residuals. Moreover, if $i = j$, only the second term survives.

$$\beta_{j,j} = -\frac{r_j^\dagger r_j}{r_{j-1}^\dagger r_{j-1}} \tag{C.8}$$

This matches the definition in Algorithm 3 and also clarifies the minus sign discrepancy in computing p_{j+1} . With this we have tied all loose ends of the conjugate gradient algorithm.

BIBLIOGRAPHY

- [1] A. Abdel-Rehim, C. Alexandrou, M. Constantinou, K. Hadjiyiannakou, K. Jansen, C. Kallidonis, G. Koutsou, and A. Vaquero Aviles-Casco (ETM), Phys. Rev. Lett. **116**, 252001 (2016), 1601.01624.
- [2] T. Métivet, Theses, Université Paris Sud - Paris XI (2015), URL <https://tel.archives-ouvertes.fr/tel-01249608>.
- [3] S. Aoki, N. Ishii, K.-I. Ishikawa, N. Ishizuka, T. Izubuchi, D. Kadoh, K. Kanaya, Y. Kuramashi, Y. Namekawa, O. H. Nguyen, et al., Progress of Theoretical and Experimental Physics **2012**, 01A102 (2012), /oup/backfile/content_public/journal/ptep/2012/1/10.1093/ptep/pts002/2/pts002.pdf URL + <http://dx.doi.org/10.1093/ptep/pts002>.
- [4] S. D. Bass and A. W. Thomas, Physics Letters B **684**, 216 (2010), ISSN 0370-2693, URL <http://www.sciencedirect.com/science/article/pii/S0370269310000377>.
- [5] C. A. Aidala, S. D. Bass, D. Hasch, and G. K. Mallot, Rev. Mod. Phys. **85**, 655 (2013), 1209.2803.
- [6] H. W. Babcock, Lick Observatory Bulletin **19**, 41 (1939).
- [7] E. Corbelli and P. Salucci, Mon. Not. Roy. Astron. Soc. **311**, 441 (2000), astro-ph/9909252.

- [8] S. M. Faber and R. E. Jackson, *The Astrophysical Journal* **204**, 668 (1976).
- [9] A. N. Taylor, S. Dye, T. J. Broadhurst, N. Benítez, and E. van Kampen, *The Astrophysical Journal* **501**, 539 (1998), URL <http://stacks.iop.org/0004-637X/501/i=2/a=539>.
- [10] J. R. Ellis, K. A. Olive, and C. Savage, *Phys. Rev.* **D77**, 065026 (2008), 0801.3656.
- [11] A. Bottino, F. Donato, N. Fornengo, and S. Scopel, *Phys. Rev.* **D78**, 083520 (2008), 0806.4099.
- [12] R. J. Hill and M. P. Solon, *Phys. Lett.* **B707**, 539 (2012), 1111.0016.
- [13] V. Bernard, N. Kaiser, and U.-G. Meissner, *Phys. Lett.* **B389**, 144 (1996), [hep-ph/9607245](http://arxiv.org/abs/hep-ph/9607245).
- [14] M. M. Pavan, I. I. Strakovsky, R. L. Workman, and R. A. Arndt, *PiN Newslett.* **16**, 110 (2002), [hep-ph/0111066](http://arxiv.org/abs/hep-ph/0111066).
- [15] J. M. Alarcon, J. Martin Camalich, and J. A. Oller, *Phys. Rev.* **D85**, 051503 (2012), 1110.3797.
- [16] B. Borasoy and U.-G. Meissner, *Annals Phys.* **254**, 192 (1997), [hep-ph/9607432](http://arxiv.org/abs/hep-ph/9607432).
- [17] R. D. Young, *PoS LATTICE2012*, 014 (2012), 1301.1765.
- [18] G. S. Bali, S. Collins, D. Richtmann, A. Schäfer, W. Söldner, and A. Sternbeck (RQCD), *Phys. Rev.* **D93**, 094504 (2016), 1603.00827.
- [19] T. Marrodán Undagoitia and L. Rauch, *J. Phys.* **G43**, 013001 (2016), 1509.08767.
- [20] J. Aubert, G. Bassompierre, K. Becks, C. Best, E. Böhm, X. de Bouard, F. Brasse, C. Broll, S. Brown, J. Carr, et al., *Physics Letters B* **123**, 275 (1983), ISSN 0370-2693, URL <http://www.sciencedirect.com/science/article/pii/0370269383904379>.

- [21] J. C. Bernauer et al. (A1), Phys. Rev. Lett. **105**, 242001 (2010), 1007.5076.
- [22] P. J. Mohr, B. N. Taylor, and D. B. Newell, Rev. Mod. Phys. **84**, 1527 (2012), URL <http://link.aps.org/doi/10.1103/RevModPhys.84.1527>.
- [23] R. Pohl, R. Gilman, G. A. Miller, and K. Pachucki, Ann. Rev. Nucl. Part. Sci. **63**, 175 (2013), 1301.0905.
- [24] R. Pohl et al. (2016), 1607.03165.
- [25] Z. Ahmed et al. (HAPPEX), Phys. Rev. Lett. **108**, 102001 (2012), 1107.0913.
- [26] J. Green, S. Meinel, M. Engelhardt, S. Krieg, J. Laeuchli, J. Negele, K. Orginos, A. Pochinsky, and S. Syritsyn, Phys. Rev. **D92**, 031501 (2015), 1505.01803.
- [27] R. S. Sufian, Y.-B. Yang, J. Liang, T. Draper, and K.-F. Liu (2017), 1705.05849.
- [28] C. E. Carlson, Prog. Part. Nucl. Phys. **82**, 59 (2015), 1502.05314.
- [29] D. S. Armstrong and R. D. McKeown, Ann. Rev. Nucl. Part. Sci. **62**, 337 (2012), 1207.5238.
- [30] M. Creutz, *Quarks, Gluons and Lattices*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, 1983), ISBN 9780521315357, URL <https://books.google.com/books?id=mcCyB3ewyeMC>.
- [31] C. Gattringer and C. Lang, *Quantum Chromodynamics on the Lattice: An Introductory Presentation*, Lecture Notes in Physics (Springer Berlin Heidelberg, 2009), ISBN 9783642018497, URL <https://books.google.com/books?id=12hZKnlYDxoC>.
- [32] J. Collins, *Foundations of Perturbative QCD*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology (Cambridge University Press, 2011), ISBN 9781139500623, URL <https://books.google.com/books?id=0xGi1KW9vykC>.

- [33] J. B. Kogut and L. Susskind, Phys. Rev. **D11**, 395 (1975).
- [34] D. B. Kaplan, Physics Letters B **288**, 342 (1992), ISSN 0370-2693, URL <http://www.sciencedirect.com/science/article/pii/037026939291112M>.
- [35] Y. Shamir, Nucl. Phys. **B406**, 90 (1993), hep-lat/9303005.
- [36] V. Furman and Y. Shamir, Nucl. Phys. **B439**, 54 (1995), hep-lat/9405004.
- [37] R. Frezzotti, P. A. Grassi, S. Sint, and P. Weisz (Alpha), JHEP **08**, 058 (2001), hep-lat/0101001.
- [38] M. Lüscher and P. Weisz, Comm. Math. Phys. **97**, 59 (1985), URL <http://projecteuclid.org/euclid.cmp/1103941978>.
- [39] B. Sheikholeslami and R. Wohlert, Nucl.Phys. **B259**, 572 (1985).
- [40] M. Luscher, S. Sint, R. Sommer, and P. Weisz, Nucl. Phys. **B478**, 365 (1996), hep-lat/9605038.
- [41] M. Luscher, S. Sint, R. Sommer, P. Weisz, H. Wittig, and U. Wolff, Nucl. Phys. Proc. Suppl. **53**, 905 (1997), hep-lat/9608049.
- [42] G. Kilcup, S. Sharpe, R. Gupta, G. Guralnik, A. Patel, and T. Warnock, Physics Letters B **164**, 347 (1985), ISSN 0370-2693, URL <http://www.sciencedirect.com/science/article/pii/0370269385903399>.
- [43] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. V. der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition* (SIAM, Philadelphia, PA, 1994).

- [44] H. A. Van der Vorst, *SIAM Journal on scientific and Statistical Computing* **13**, 631 (1992).
- [45] Y. Saad, *Iterative Methods for Sparse Linear Systems* (Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2003), 2nd ed., ISBN 0898715342.
- [46] R. Babich, J. Brannick, R. C. Brower, M. A. Clark, T. A. Manteuffel, S. F. McCormick, J. C. Osborn, and C. Rebbi, *Phys. Rev. Lett.* **105**, 201602 (2010), 1005.3043.
- [47] A. Frommer, K. Kahl, S. Krieg, B. Leder, and M. Rottmann, *SIAM J. Sci. Comput.* **36**, A1581 (2014), 1303.1377.
- [48] M. A. Clark, B. Joó, A. Strelchenko, M. Cheng, A. Gambhir, and R. C. Brower, in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis* (IEEE Press, Piscataway, NJ, USA, 2016), SC '16, pp. 68:1–68:12, ISBN 978-1-4673-8815-3, URL <http://dl.acm.org/citation.cfm?id=3014904.3014995>.
- [49] T. Iitaka and T. Ebisuzaki, *Phys. Rev. E* **69**, 057701–1–057701–4 (2004).
- [50] M. N. Wong, F. J. Hickernell, and K. I. Liu, *Tech. Rep. 377(7/04)*, Hong Kong Baptist University (2004).
- [51] H. Avron and S. Toledo, *Journal of the ACM* **58**, Article 8 (2011).
- [52] J. Tang and Y. Saad, *Report UMSI 2010/114* (????).
- [53] R. Babich, R. Brower, M. Clark, G. Fleming, J. Osborn, C. Rebbi, and D. Schaich (4 May 2011), 1012.0562v2.

- [54] C. Morningstar, J. Bulava, J. Foley, K. Juge, D. Lenkner, M. Peardon, and C. Wong1, Phys. Rev. D **83** (2011), 1104.3870v1.
- [55] A. Stathopoulos, J. Laeuchli, and K. Orginos (2013), 1302.4018.
- [56] A. S. Gambhir, A. Stathopoulos, and K. Orginos (2016), 1603.05988.
- [57] T. A. Davis and Y. Hu, ACM Trans. Math. Softw. **38**, 1:1 (2011), ISSN 0098-3500, URL <http://doi.acm.org/10.1145/2049662.2049663>.
- [58] R. Morgan and W. Wilcox, Tech. Rep. BU-HEPP-04-01, Baylor University (2004).
- [59] A. Stathopoulos and K. Orginos, SIAM J. Sci. Comput. **32**, 439 (2010), 0707.0131.
- [60] A. Abdel-Rehim, K. Orginos, and A. Stathopoulos, PoS **LAT2009**, 036 (2009), 0911.2285.
- [61] A. Stathopoulos and J. R. McCombs, *Primme: Preconditioned iterative multimethod eigensolver: Methods and software description* (2006).
- [62] B. Yoon et al. (2016), 1602.07737.
- [63] B. Yoon et al., Phys. Rev. **D95**, 074508 (2017), 1611.07452.
- [64] S. Duane, A. Kennedy, B. J. Pendleton, and D. Roweth, Physics Letters B **195**, 216 (1987), ISSN 0370-2693, URL <http://www.sciencedirect.com/science/article/pii/037026938791197X>.
- [65] S. Borsanyi et al., JHEP **09**, 010 (2012), 1203.4469.
- [66] G. S. Bali, S. Collins, and A. Schaefer (2010), 0910.3970v2.
- [67] T. Blum, T. Izubuchi, and E. Shintani, Phys. Rev. **D88**, 094503 (2013), 1208.4349.

- [68] M. Foster and C. Michael (UKQCD), Phys. Rev. **D59**, 074503 (1999), [hep-lat/9810021](#).
- [69] C. Michael, M. S. Foster, and C. McNeile (UKQCD), Nucl. Phys. Proc. Suppl. **83**, 185 (2000), [hep-lat/9909036](#).
- [70] C. Alexandrou et al. (2017), [1703.08788](#).
- [71] S. Syritsyn, A. S. Gambhir, B. Musch, and K. Orginos, PoS **LATTICE2016**, 176 (2017).
- [72] G. Martinelli, C. Pittori, C. T. Sachrajda, M. Testa, and A. Vladikas, Nucl. Phys. **B445**, 81 (1995), [hep-lat/9411010](#).
- [73] C. Sturm, Y. Aoki, N. H. Christ, T. Izubuchi, C. T. C. Sachrajda, and A. Soni, Phys. Rev. **D80**, 014501 (2009), [0901.2599](#).
- [74] J. Green, N. Hasan, S. Meinel, M. Engelhardt, S. Krieg, J. Laeuchli, J. Negele, K. Orginos, A. Pochinsky, and S. Syritsyn (2017), [1703.06703](#).
- [75] C. Michael, C. McNeile, and D. Hepburn (UKQCD), Nucl. Phys. Proc. Suppl. **106**, 293 (2002), [hep-lat/0109028](#).
- [76] T. Bhattacharya, R. Gupta, W. Lee, S. R. Sharpe, and J. M. S. Wu, Phys. Rev. **D73**, 034504 (2006), [hep-lat/0511014](#).
- [77] M. J. Savage, P. E. Shanahan, B. C. Tiburzi, M. L. Wagman, F. Winter, S. R. Beane, E. Chang, Z. Davoudi, W. Detmold, and K. Orginos, PoS **ICHEP2016**, 506 (2016), [1611.00344](#).

- [78] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*, Advanced book classics (Addison-Wesley Publishing Company, 1995), ISBN 9780201503975, URL <https://books.google.com/books?id=i35LALN0GosC>.

ARJUN SINGH GAMBHIR

3901 Railroad Ave, Fairfax, 22030 | 571-338-3564 | asgambhir@email.wm.edu

EDUCATION

College of William and Mary

PhD in Physics

MS in Physics

Virginia Polytechnic Institute and State University

BS in Physics, Minor in Astronomy

Williamsburg, VA

June 2017

December 2013

Blacksburg, VA

May 2011

TEACHING EXPERIENCE

College of William and Mary

Physics for Life Sciences Instructor (Faculty)

- Taught a full summer class on second semester introductory physics and held office hours.
- Syllabus included thermodynamics, basic electricity and magnetism, geometric and wave optics, and radioactivity.

Lab Teaching Assistant

- Taught 2 sections per semester of introductory physics labs, graded, and held office hours.
- Discussed intro topics: conservation of energy, momentum, torque, sound waves.

Georgetown University

Intermediate Electricity and Magnetism Teaching Assistant

- Drafted homework solutions, graded, and arranged weekly office hours for E&M course.
- Developed supplemental material for recitations.

Tutorial/Recitation Teaching Assistant

- Ran recitation sections consisting of lecturing, problem help sessions, and in class worksheets for premedical physics. Held 8 or more office hours most weeks.
- Implemented the McDermott *et al* Tutorials for the first time in the context of premedical physics at Georgetown.

Williamsburg, VA

July 2014-August 2014

Aug 2012-December 2013

Washington, D.C.

January 2012-May 2012

Jun 2011-May 2012

RESEARCH SKILLS/EXPERIENCE

Current Research Interests

- Nucleon and nuclear structure, especially disconnected contributions from sea quarks.
- Improved algorithms for Monte Carlo techniques useful in Lattice QCD computation.
- Numerical Linear Algebra, in particular: inverting or finding trace functions of incredibly large, sparse matrices.

College of William and Mary

Thesis Research

- Developed a new class of algorithms with the potential to alleviate critical slowing down in Lattice QCD.
- Studied the effects of singular value trace deflation in a variety of fields, including Signal Processing, Computational Fluid Dynamics, and Lattice QCD. For relevant problems in QCD (disconnected diagrams), a factor of 60 speedup was observed.
- Helped with testing, interfacing, and deployment of GPU multigrid solver using fine-grained parallelism for Lattice Field Theory calculations.
- Applied deflation algorithms to compute disconnected contributions of nucleon matrix elements and form factors.

Georgetown University

Graduate Research

- Analyzed results of Summer McDermott *et al* closed Pre-tests and Post-tests to develop a before/after snapshot of student performance. Executed a variety of teaching techniques and cross referenced with corresponding data evaluation.
- Delivered FMCE to fall population, and BEMA to spring semester students. Steered exploration of comparison to previous non-tutorial years.

Virginia Polytechnic Institute and State University

Undergraduate Research

- Conducted research with Electrochromism to explore how properties of Electrochromic materials affect speed and contrast in color switching.
- Converted Indium-Tin Oxide (ITO) coated glass slide into Electrochromic film using Layer-by-Layer Deposition method. Trained to use Bi-Layer Dipping Machine, Spectrometer and various Thickness Measuring Instruments.

Williamsburg, VA

June 2012 – Present

Washington, D.C.

Jun 2011 – May 2012

Blacksburg, VA

Aug 2010 – May 2011

RECENT PRESENTATIONS AND WORKSHOPS

Lawrence Livermore National Lab

Livermore, CA
December 2016

Computing Disconnected Diagrams in Lattice QCD

- Gave a seminar on motivations and strategies for computing disconnected diagrams in the ab initio theory of quark and gluon interactions.
- Presented state of the art methods developed from the William and Mary Lattice group which demonstrate over an order of magnitude variance reduction compared to standard techniques for computing disconnected observables.

Oak Ridge National Lab 2016 GPU Hackathon

Knoxville, TN
October 2016

Team Hybrid Titans

- Participated in a professional hackathon to extend functionality of production-level research software.
- Engineered a solution to large multigrid setup cost by persisting the multigrid subspace in a Named Object Buffer to allow preconditioner to be used through multiple trajectories of Hybrid Monte Carlo.

The 34th International Symposium on Lattice Field Theory

Southampton, U.K.
July 2016

Algorithms for Disconnected Diagrams

- Delivered a talk focusing on the synergy of two algorithms: Hierarchical Probing and Deflation.
- Explored this synergy theoretically and experimentally, presented numerical results of disconnected nucleon charges that demonstrated substantial cost reduction.

Thomas Jefferson National Accelerator Facility Open House

Newport News, VA
April 2016

The Glue in You and Everything You Love

- Spoke about effects from the Higgs on the mass of nucleons and QCD dynamics.
- Presented recent Lattice QCD results of hadron spectrum at varying quark masses.

The 32nd International Symposium on Lattice Field Theory

Ney York, NY
June 2014

Algorithms that Alleviate Critical Slow Down in Lattice QCD

- Presented work on a new class of algorithms with the potential to alleviate critical slowing down when approaching the continuum limit in a lattice field theory.

References per request