



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Non-Perturbative Renormalization and Low Mode Averaging with Domain Wall Fermions



Rudy Arthur

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
to the
University of Edinburgh
April 2012

Abstract

This thesis presents an improved method to calculate renormalization constants in a regularization invariant momentum scheme using twisted boundary conditions. This enables us to simulate with momenta of arbitrary magnitude and a fixed direction. With this new technique, together with non-exceptional kinematics and volume sources, we are able to take a statistically and theoretically precise continuum limit. Thereafter, all the running of the operators with momentum scale is due to their anomalous dimension. We use this to develop a practical scheme for step scaling with off shell vertex functions. We develop the method on $16^3 \times 32$ lattices to show the practicality of using small volume simulations to step scale to high momenta. We also use larger $24^3 \times 64$ and $32^3 \times 64$ lattices to compute renormalization constants very accurately. Combining these with previous analyses we are able to extract a precise value for the light and strange quark masses and the neutral kaon mixing parameter B_K .

We also analyse eigenvectors of the domain wall Dirac matrix. We develop a practical and cost effective way to compute eigenvectors using the implicitly restarted Lanczos method with Chebyshev acceleration. We show that calculating eigenvectors to accelerate propagator inversions is cost effective when as few as one or two propagators are required. We investigate the technique of low mode averaging (LMA) with eigenvectors of the domain wall matrix for the first time. We find that for low energy correlators, pions for example, LMA is very effective at reducing the statistical noise. We also calculated the η and η' meson masses, which required evaluating disconnected correlation functions and combining stochastic sources with LMA.

Declaration

Except where otherwise stated, the research undertaken in this thesis was my unaided work, performed as a member of the RBC-UKQCD collaboration. This work has not been submitted for any other award or professional qualification.

This work used gauge configurations generated by the RBC-UKQCD collaboration. Twisted vertex functions, see chapter 4, for the 16^3 $\beta = 2.13$ and $\beta = 2.23$ lattices were generated by me, as well as some of the twisted renormalization data on the 24^3 and 32^3 lattices. The analysis code was written using UKHadron, a code base principally developed by Peter Boyle. Some of the analysis was cross-checked by Dirk Brommel, Z_{BK} , and Nicolas Garron, Z_m .

The eigenvector code, chapter 5, was written by me using the bagel and bfm libraries developed by Peter Boyle for the multiplication operators. The contraction code for chapter 6 was also written by me using UKHadron and all eigenvector calculations, quark contractions and analyses were performed by me.

Some results presented in this thesis appear in,

- **Step Scaling with off-shell renormalisation. Phys.Rev. D83 (2011) 114511**
- **Step Scaling with off-shell renormalisation. PoS LATTICE2010 (2010) 244**

These describe Rome-Southampton renormalization with twisted boundary conditions and detail our step scaling procedure. I performed an analysis of two 16^3 lattices. This involved generating vertex functions, measuring the lattice spacing on the $\beta = 2.23$ lattice using Wilson loops and performing the analysis of the renormalization constants and step scaling functions, this was verified by Peter Boyle. The second work contains an analysis of finite volume effects on non-perturbative vertex functions.

-
- **Continuum Limit of B_K from 2+1 Flavor Domain Wall QCD.** Phys.Rev. D84 (2011) 014503

Data for the renormalization of B_K was generated by me and I performed the analysis of the renormalization constants. This was verified by Dirk Brommel.

- **Continuum Limit Physics from 2+1 Flavor Domain Wall QCD.** Phys.Rev. D83 (2011) 074508

I cross checked the Z_m analysis and generated data for the measurement of Z_V .

I have also contributed to other published works not described in this thesis:

- **Opening the Rome-Southampton window for operator mixing matrices.** Phys.Rev. D85 (2012) 014501
- **Lattice Results for Low Moments of Light Meson Distribution Amplitudes.** Phys.Rev. D83 (2011) 074505
- **Light Meson Distribution Amplitudes.** PoS LATTICE2010 (2010) 154

R. Arthur
2012

Acknowledgements

Firstly I would like to thank my supervisor, Peter Boyle, for his support, dedication and energy over the last few years. My PhD was funded by SUPA, for which I remain grateful.

I am also grateful to all the people who helped me over the course of my PhD including Luigi Del Debbio, Nicolas Garron, Jamie Hudspith, Chris Kelly, Tony Kennedy, James Zanotti, Taku Izubuchi, Eigo Shintani, Dirk Brommel, Jonathan Flynn, Tom Rae and Ben Samways.

Thanks Leigh, Dad and especially Mam for listening to me complain about this thesis. Thanks to Arwen for sending me chocolate. You all helped more than you realise.

Contents

Abstract	i
Declaration	ii
Acknowledgements	iv
Contents	v
1 Lattice QCD	1
1.1 Lattice QCD and the Standard Model	1
1.2 Path Integrals	2
1.3 Basic Lattice Formulation	5
1.4 Domain Wall Fermions	10
1.4.1 Chiral Symmetry	14
1.5 Overlap Fermions	17
1.6 Equivalence of Overlap and Domain Wall	18
1.7 Pauli-Villars Fields	23
2 Renormalization	25
2.1 Wilson's Renormalization Group	25
2.1.1 ϕ^4 theory	29
2.2 QCD and the continuum limit	32
2.3 Callan Symanzik Equation and Scaling	36
2.4 Composite Operator Renormalization	38
2.5 $\overline{MS}$	39
3 Lattice Methods	42
3.1 Correlation Functions	42
3.1.1 Mesons	44
3.1.2 Disconnected Diagrams	45
3.1.3 Four-Quark Operators	45
3.2 Dirac Inversion	47
3.2.1 Preconditioning	48
3.3 Gauge Fixing	50

3.4	Fitting	51
3.4.1	Black Box Fits	53
3.5	Sommer Scale	54
4	Step Scaling with Rome-Southampton Renormalization	57
4.1	Rome-Southampton Renormalization	58
4.1.1	RI-MOM	58
4.1.2	RI-SMOM	63
4.1.3	Matching	66
4.2	Momentum Sources	68
4.3	Twisted Boundary Conditions	70
4.4	Four Quark Operators and B_K	74
4.5	Step Scaling	76
4.5.1	Previous Work on Step Scaling	77
4.5.2	RI-MOM Step Scaling Method	78
4.6	Results on Small Lattices	84
4.6.1	Vertex Functions	84
4.6.2	Scale determination in a small volume	96
4.6.3	Scale Evolution Functions	99
4.7	Volume Dependence	100
4.8	Results on Large Lattices	105
4.8.1	Bilinear Renormalization	107
4.8.2	B_K Renormalization	121
4.9	Conclusion	123
5	Eigenvectors of the Dirac Operator	130
5.1	Algorithms	131
5.1.1	Arnoldi Algorithm	133
5.1.2	Lanczos Algorithm	135
5.2	Implementation	137
5.2.1	Implicit Restarting	137
5.2.2	Locking	140
5.2.3	Gram Schmidt	140
5.2.4	Chebyshev filters	142
5.3	Algorithm Tests	143
5.3.1	Tuning	146
5.4	Deflation	153
5.4.1	Deflation: Results	155
5.5	Instantons	161
5.6	Conclusions	166

6	Low Mode Averaging	167
6.1	Low Mode Averaging with Domain Wall Fermions	167
6.1.1	Effective Overlap Operator	169
6.1.2	Hermitian DWF Operator	170
6.1.3	Preconditioned DWF	171
6.2	Estimating the High Mode Part	173
6.2.1	Point Sources	173
6.3	Stochastic Sources	175
6.3.1	Connected Correlation Functions	175
6.3.2	Disconnected Correlation Functions	177
6.3.3	LMA with Stochastic Sources	178
6.4	Meson Correlation functions	179
6.4.1	Comparison of Eigenvectors	179
6.4.2	Mass Dependence	185
6.4.3	η Meson	188
6.4.4	Results	193
6.5	Conclusions	199
	Bibliography	212
A	Eigenvalues in the Unit Gauge	213
A.1	D_W and H_W	213
A.2	H_T	215
A.3	D_{ov} and H_{ov}	216

Chapter 1

Lattice QCD

1.1 Lattice QCD and the Standard Model

The standard model of particle physics has stood since around 1974 and describes how fundamental particles interact via the electroweak and strong forces. The strong force, described by QCD, is asymptotically free which means that at high energies the strength of the force is reduced and the coupling constant is small. Small coupling means perturbation theory works. However at low energies, $O(1\text{GeV})$, the coupling is large, quarks are strongly bound and perturbation theory fails. Lattice QCD enables us to study the strong force in this regime.

The rest of this chapter describes how to formulate QCD on a spacetime lattice. Briefly, we simulate the QCD vacuum which contains gluon fields and virtual quark anti-quark pairs. In this background measurements of QCD observables are made and experimentally testable predictions, such as matrix elements and hadron masses, are extracted. Lattice QCD is extremely computationally intensive which initially limited its usefulness, however, after 30 years of progress in high performance computing and algorithms high precision lattice QCD is here. Many phenomenologically important quantities can now be calculated with errors of the order of a few percent or less.

This thesis follows two tracks. Given the ability to perform high precision calculations with lattice QCD, theoretical uncertainties need to be brought under control. One of the great successes of lattice QCD has been in the prediction of the neutral kaon mixing parameter B_K which helps constrain CP violation in the standard model. We now find that the accuracy of this prediction is

limited by perturbation theory. B_K , and many other quantities measured via lattice QCD, must be renormalized. The method of choice for this, for the RBC-UKQCD collaboration and also for many other groups, is Rome-Southampton renormalization. Chapter 4 details how we and others have improved the original formulation. The Rome-Southampton method is very useful for combining lattice QCD with perturbative predictions. However, lattice calculations are done at low energy, where perturbation theory is inaccurate and matching at low energy can introduce large uncertainties into otherwise precise lattice calculations. Chapter 4 also describes a method to mitigate this.

The second track is concerned with continuing the march towards ever higher statistical precision. Measurements of QCD observables containing quarks require us to calculate quark propagators $S(x, y)$; the probability amplitude for a quark to propagate from x to y . Typically there are too many points on the lattice to calculate this between every x and y and instead we calculate the propagator from one or at most a few points, x , to every y . This means we are not extracting all possible information from the lattice about how the quarks propagate in a given background, which leads to greater statistical uncertainty. Chapters 5 and 6 describe a method to try and extract more information from each lattice and thus reduce statistical uncertainties.

These two tracks represent a way for lattice QCD to become an even more important tool in particle physics; one that heavily constrains searches for physics beyond the standard model and gives extremely accurate predictions of quark physics at low energy. We now turn to QCD and its formulation on the lattice.

1.2 Path Integrals

Lattice QCD is based on the path integral formulation of quantum field theory. In Euclidean space the generating functional W is defined by a path integral,

$$W[J, \bar{\sigma}, \sigma] = \int \mathcal{D}A \mathcal{D}\psi \mathcal{D}\bar{\psi} \exp \left(-S[\psi, \bar{\psi}, A] + \int d^4x \bar{\sigma}(x) \psi(x) + \sigma(x) \bar{\psi}(x) + J_\mu A^\mu \right) \quad (1.1)$$

This is the generating functional of a theory containing fermions, ψ , and gauge fields, A . The total action is S and we include sources J_μ for the gauge fields and (Grassmann) sources $\sigma, \bar{\sigma}$ for the fermions. In lattice field theory the

integral is made well-defined and finite by discretizing spacetime and simulating in a finite volume, giving a finite but very large dimensional integral. Monte Carlo techniques then allow us to calculate correlation functions and from these, physical quantities, such as masses and form factors. Since this is not a perturbative method it does not rely on a small coupling expansion so that, in the low energy regime of QCD where the coupling constant is large, see chapter 2, lattice QCD works while perturbation theory fails. The only approximation is that we have a finite lattice spacing; provided this is small we can extrapolate it to zero in a rigorous way. This allows us to calculate low energy properties of QCD without introducing models or phenomenological parameters. In exchange we have to deal with discretization effects and statistical noise from the Monte Carlo.

The action for QCD separates into pure gauge and quark sectors,

$$S = S_G[A] + \bar{\psi} D[A] \psi, \quad (1.2)$$

where the quark fields are anti-commuting Grassmann variables. Representing these on a computer is difficult and they are instead integrated out using the standard formula for the integral of a Gaussian of Grassmann fields,

$$\int \mathcal{D}\psi \mathcal{D}\bar{\psi} e^{-\bar{\psi} D[A] \psi} = \det(D[A]). \quad (1.3)$$

$D(A)$ is a matrix of ordinary complex numbers and its determinant can be found in principle. Numerical integration of extremely large dimensional integrals is almost always done by Monte Carlo. For lattice QCD naive Monte Carlo will fail because of the vast number of possible configurations to sample. What is done in practice is importance sampling. A few, perhaps $O(10^3)$, configurations that give large contributions to the partition function are generated by a Markov process with probability distribution,

$$P(A) = \frac{\det(D[A]) e^{-S_G[A]}}{W} \quad (1.4)$$

and then used to evaluate the integral. Expectation values,

$$\langle \mathcal{O}(A) \rangle \equiv \langle 0 | T [\mathcal{O}(A)] | 0 \rangle = \frac{1}{W} \int \mathcal{D}A \det(D[A]) e^{-S_G[A]} \mathcal{O}(A) \quad (1.5)$$

(the sources are set to zero here) are then evaluated via

$$\langle \mathcal{O}(A) \rangle = \frac{1}{N} \sum_i^N \mathcal{O}(A_i). \quad (1.6)$$

Quark fields in $\mathcal{O}$ are contracted together according to Wick's theorem. $\mathcal{O}$ will be some number of fermion propagators, $S[A] = D^{-1}[A]$, with linear operators $(\partial_\mu, \gamma_\mu, \tau_a, \dots)$ acting on them, eg. if $\mathcal{O}(x, y) = \psi(x)\bar{\psi}(y)$ then $\langle \mathcal{O}(x, y) \rangle$ is the probability amplitude for a quark to propagate from y to x .

Evaluating the determinant, $\det(D[A])$, is challenging. $D[A]$ is a $12V \times 12V$ matrix with V , the lattice volume, of order $O(10^6)$ and it is not possible to evaluate its determinant directly, eg. by eigenvalues or some other simple method. Instead, one exploits the relative inverse between Gaussian integrals of Grassmann and c-number variables. For ordinary complex valued fields ϕ ,

$$\int \mathcal{D}\phi \mathcal{D}\phi^\dagger e^{-\phi^\dagger D^{-1}[A] \phi} = \det(D[A]), \quad (1.7)$$

gives the determinant of D . Introducing scalar fields ϕ , called pseudofermions [135], we can write the partition function as

$$Z = \int \mathcal{D}A \mathcal{D}\phi \mathcal{D}\phi^\dagger e^{-S_G[A] - \phi^\dagger D^{-1}[A] \phi}. \quad (1.8)$$

The sources are set to zero and now everything is written in terms of complex valued fields that can be manipulated by a computer. In the past the determinant was set to a constant, called the quenched approximation, to make the computation cheaper. This amounts to neglecting quark loops and is an uncontrolled approximation, [43]. Today, and in all the simulations that will be considered in this thesis, computers and algorithms have advanced far enough that including the determinant is possible, at least for unphysically large quark masses, though modern simulations are rapidly approaching the physical point. Generating configurations according to the probability distribution of equation 1.4 is a very difficult task and beyond our scope. Most of the configurations used in this thesis were generated by the RBC-UKQCD collaboration with the rational hybrid Monte Carlo algorithm [40], [41], a modification of hybrid Monte Carlo, [54].

1.3 Basic Lattice Formulation

When regularizing a quantum field theory we perforce break one or more of its symmetries. In Euclidean space going over to a lattice breaks the rotational invariance $\mathcal{O}(4)$ down to $\mathcal{H}(4)$, the group of symmetries of a $4D$ hypercube. The infinite volume and continuum limits restore this symmetry in a straightforward way. The chiral symmetry of massless QCD is more subtle. Quantum effects break this symmetry even in the continuum. The way to deal with this on the lattice is described in the following sections and more detail can be found in textbooks on lattice QCD; for example references [70], [117] and [105].

The continuum, Euclidean action for fermions coupled to gauge fields is,

$$S_F = \int d^4x \bar{\psi}(x) (\gamma_\mu (\partial_\mu + iA_\mu) + m) \psi(x) \quad (1.9)$$

with the equations of motion given by,

$$(\gamma_\mu (\partial_\mu + iA_\mu) + m) \psi(x) = D\psi(x) = 0, \quad (1.10)$$

the gauge field is $A_\mu = t^a A_\mu^a$, with $t^a = \frac{\lambda^a}{2}$ the basis elements of the Lie algebra of $SU(3)$ and the λ^a are Gell-Mann matrices. We have also made the field redefinition $A_\mu \rightarrow \frac{A_\mu}{g}$ to remove the coupling constant, g , from the covariant derivative, it will reappear later multiplying the pure gauge term in the action.

To discretize the path integral first consider the free field case with no gauge fields. Using the symmetric discretization of the partial derivative

$$\partial_\mu(x, y) \rightarrow \frac{1}{2a} (\delta(x + \hat{\mu}, y) - \delta(x - \hat{\mu}, y)) \quad (1.11)$$

leads to the action,

$$S_F^{naive} = \sum_{x,y} \bar{\psi}(x) D_N(x, y) \psi(y) \quad (1.12)$$

with the “naive” Dirac operator given by

$$D_N(x, y) = m\delta(x, y) + \frac{1}{2a} \sum_{\mu} \gamma_\mu (\delta(x + \hat{\mu}, y) - \delta(x - \hat{\mu}, y)), \quad (1.13)$$

the spinor indices have been suppressed.

QCD is invariant under local $SU(3)$ transformations of the fermion fields,

$$\begin{aligned}\psi_a(x) &\rightarrow G_{ab}(x)\psi_b(x) \\ \bar{\psi}_a(x) &\rightarrow \bar{\psi}_b(x)G_{ba}^{-1}(x)\end{aligned}\tag{1.14}$$

where $G(x)$ is an element of the $SU(3)$ (colour) group, the group indices will also be suppressed from now on, to save cluttering the formulas. The derivatives in the Dirac operator couple adjacent sites which means the action contains gauge non-invariant terms like $\bar{\psi}(x)\psi(x+\hat{\mu})$. The gauge link couples together adjacent sites in a way that cancels the transformations of the fermion fields. Define,

$$U(x, y) = P[e^{ig \int_x^y dz_\mu A_\mu(z)}]\tag{1.15}$$

where A_μ is the usual gauge field and $P[\dots]$ is the path ordering symbol. Under a gauge transformation

$$A_\mu^a \rightarrow A_\mu^a - \frac{1}{g}\partial_\mu \alpha^a - f^{abc}\alpha^b A_\mu^c\tag{1.16}$$

the link $U(x, y)$ transforms as

$$U(x, y) \rightarrow G(x)U(x, y)G^{-1}(y).\tag{1.17}$$

with $G(x) = e^{i\alpha(x)} = e^{i\alpha^a(x)t^a}$. Inserting a gauge link to connect adjacent sites in the action we get the discretised covariant derivative,

$$D_\mu \rightarrow \frac{1}{2a} (U_\mu(x)\delta(x+\hat{\mu}, y) - U_{-\mu}(x)\delta(x-\hat{\mu}, y))\tag{1.18}$$

where $U_\mu(x) = U(x, x+\hat{\mu})$ and $U_{-\mu}(x) = U_\mu^\dagger(x-\hat{\mu})$. Inserting covariant derivatives changes the naive Dirac action and operator to,

$$S_F^{naive} = \sum_{x,y} \bar{\psi}(x)D_N(x, y)\psi(y)\tag{1.19}$$

$$\begin{aligned}D_N(x, y) &= m\delta(x, y) + \\ &\quad \frac{1}{2a} \sum_\mu \gamma_\mu (U_\mu(x)\delta(x+\hat{\mu}, y) - U_{-\mu}(x)\delta(x-\hat{\mu}, y))\end{aligned}\tag{1.20}$$

This simple fermion discretisation introduces unphysical behaviour: 15 extra

fermion species appear in the spectrum. On Fourier transforming $D(x, y)$ in the free field, $U_\mu(x) = 1$, we get,

$$\begin{aligned} D(p, q) &= \frac{1}{V} \sum_{x, y} e^{ipx} D(x, y) e^{-iqy} \\ &= \delta(p, q) \tilde{D}(p) \end{aligned}$$

$$\tilde{D}(p) = m + \frac{i}{a} \sum_{\mu} \gamma_{\mu} \sin(p_{\mu} a) \quad (1.21)$$

$$\tilde{D}^{-1}(p) = \frac{m - \frac{i}{a} \sum_{\mu} \gamma_{\mu} \sin(p_{\mu} a)}{m^2 + \frac{1}{a^2} \sum_{\mu} \sin^2(p_{\mu} a)}. \quad (1.22)$$

As well as a pole at $p = (0, 0, 0, 0)$ there is a pole at each of the other 15 corners of the Brioullin zone, $p_{\mu} = \{0, \frac{\pi}{a}\}$. These poles will appear in lattice simulations as extra, mass-degenerate fermions. In order to remove these we have to modify the above formulation.

A way to remove these unphysical modes was suggested by Wilson [136]. A gauge invariant discretisation of the Laplacian, $(-\frac{a}{2})\partial_{\mu}\partial_{\mu}$, is added to the action. This has to have an extra factor of the lattice spacing, a , to be dimensionally correct and therefore vanishes in the naive continuum limit. The lattice Laplacian is,

$$-a \sum_{\mu} \frac{U_{\mu}(x)\delta(x + \hat{\mu}, y) - 2\delta(x, y) + U_{-\mu}(x)\delta(x - \hat{\mu}, y)}{2a^2}. \quad (1.23)$$

This modifies the Dirac operator to

$$D_W(x, y) = (\frac{4}{a} + m)\delta(x, y) - \frac{1}{2a} \sum_{\pm\mu} (1 - \gamma_{\mu}) U_{\mu}(x) \delta(x + \hat{\mu}, y) \quad (1.24)$$

using the convention that $\gamma_{-\mu} = -\gamma_{\mu}$. In momentum space D_W becomes,

$$\tilde{D}_W(p) = m + \frac{i}{a} \sum_{\mu} \gamma_{\mu} \sin(p_{\mu} a) + \frac{1}{a} \sum_{\mu} (1 - \cos(p_{\mu} a)). \quad (1.25)$$

Now the $p = (0, 0, 0, 0)$ pole gives the correct mass while the doublers get an additional factor $\frac{2}{a}$ for every factor of $\frac{\pi}{a}$. This gives them masses of the order of a^{-1} and they become infinitely massive in the continuum limit.

The Wilson-Dirac operator is no longer anti-Hermitian, however it still has

an important symmetry,

$$\gamma_5 D_W \gamma_5 = D_W^\dagger. \quad (1.26)$$

To demonstrate this we compare $\gamma_5 D_W \gamma_5$ and $D_W^\dagger$,

$$\gamma_5 D_W \gamma_5(x, y) = (4 + m)\delta(x, y) - \frac{1}{2a} \sum_{\pm\mu} (1 + \gamma_\mu) U_\mu(x) \delta(x + \hat{\mu}, y) \quad (1.27)$$

$$D_W^\dagger(x, y) = (4 + m)\delta(x, y) - \frac{1}{2a} \sum_{\pm\mu} (1 - \gamma_\mu) U_\mu^\dagger(y) \delta(y + \hat{\mu}, x) \quad (1.28)$$

$$\begin{aligned} \sum_{\pm\mu} (1 - \gamma_\mu) U_\mu^\dagger(y) \delta(y + \hat{\mu}, x) &= \sum_{\pm\mu} (1 - \gamma_\mu) U_\mu^\dagger(x - \hat{\mu}) \delta(x - \hat{\mu}, y) \\ &= \sum_{\pm\mu} (1 + \gamma_\mu) U_{-\mu}^\dagger(x + \hat{\mu}) \delta(x + \hat{\mu}, y) \\ &= \sum_{\pm\mu} (1 + \gamma_\mu) U_\mu(x) \delta(x + \hat{\mu}, y) \end{aligned}$$

which implies $\gamma_5 D_W \gamma_5 = D_W^\dagger$ and that,

$$H_W = \gamma_5 D_W \quad (1.29)$$

is Hermitian.

This symmetry implies that,

$$\det(D_W - \lambda I) = \det(\gamma_5 D_W \gamma_5 - \lambda I) = \det(D_W^\dagger - \lambda I). \quad (1.30)$$

Complex conjugating both sides tells us that whenever λ is an eigenvalue so is $\lambda^\dagger$. It also implies that if v is a right eigenvector, $D_W v = \lambda v$, then, since

$$(\gamma_5 v)^\dagger D_W = v^\dagger D_W^\dagger \gamma_5 = (\gamma_5 v)^\dagger \lambda^\dagger,$$

$(\gamma_5 v)^\dagger$ is a left eigenvector with eigenvalue $\lambda^\dagger$. In the free field limit the spectrum of D_W may be obtained by Fourier transforming (see appendix A for details) and is given by,

$$\lambda(p) = m + \sum_{\mu} (1 - \cos(p_\mu)) \pm i \sqrt{\sum_{\mu} \sin^2(p_\mu)} \quad (1.31)$$

which is plotted in figure 1.1 alongside eigenvalues of the free, naive Dirac operator. For the Wilson operator we see zero eigenvalues at $\{0, 2, 4, 6, 8\}$ of multiplicity $\{1, 4, 6, 4, 1\}$, corresponding to the number of factors of $\frac{\pi}{a}$. The naive operator has an eigenvalue at zero with degeneracy 16. Adding the Wilson term

has stretched out the spectrum, giving the doubler modes large real parts. Close to the origin, ie. at small momenta, the Wilson spectrum approximates the naive spectrum but with only a single pole at zero.

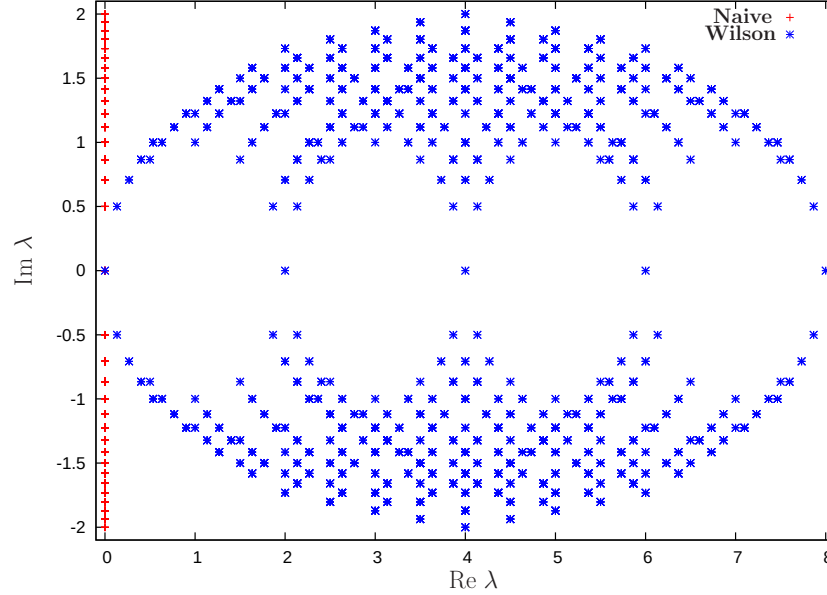


Figure 1.1: The spectrum of the free, massless Dirac operator in the naive formulation and with the Wilson term added on a 12^4 lattice.

Before discussing further improvements to the fermion formulation we first add in the action for the gauge field. The continuum Euclidean gauge action is,

$$S_G = \frac{1}{4g^2} \text{Tr} \left[\int d^4x F_{\mu\nu}(x) F_{\mu\nu}(x) \right] \quad (1.32)$$

with $F_{\mu\nu}(x) = [D_\mu, D_\nu] = \frac{\lambda_a}{2} F_{\mu\nu}^a$ and

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - f^{abc} A_\mu^b A_\nu^c. \quad (1.33)$$

f^{abc} are the structure constants of the Lie algebra of $SU(3)$. Any closed loop of gauge links is gauge invariant, the simplest being the plaquette,

$$U_{\mu\nu}(x) = U_\mu(x) U_\nu(x + \hat{\mu}) U_\mu^\dagger(x + \hat{\nu}) U_\nu^\dagger(x). \quad (1.34)$$

Using equation 1.15 this product is,

$$U_{\mu\nu}(x) = e^{igF_{\mu\nu} + \mathcal{O}(a^3)} \quad (1.35)$$

and therefore

$$\beta T r \sum_x \sum_{\mu < \nu} \left(1 - \frac{1}{2} (U_{\mu\nu}(x) + U_{\mu\nu}^\dagger(x)) \right) \rightarrow \frac{1}{4g^2} \sum_x \sum_{\mu, \nu} F_{\mu\nu}^a(x) F_{\mu\nu}^a(x). \quad (1.36)$$

in the continuum limit if we choose $\beta = \frac{6}{g^2}$ to reproduce the usual normalization of the gauge action.

This choice is not unique and there are many ways to obtain the same continuum physics. One can use this ambiguity to try to reduce or remove discretization effects at finite lattice spacing by cancelling higher order terms. Usually one considers the next most complicated lattice gauge action,

$$\beta T r \sum_x \left((1 - c_1) \sum_{\mu < \nu} \frac{1}{2} (U_{\mu\nu}(x) + U_{\mu\nu}^\dagger(x)) + c_1 \sum_{\mu \neq \nu} \frac{1}{2} (R_{\mu\nu}(x) + R_{\mu\nu}^\dagger(x)) \right). \quad (1.37)$$

$R_{\mu\nu}(x)$ is a 2×1 rectangle and the sum is over all possible orientations. By tuning c_1 perturbatively (Luscher-Weisz [98], [97]) or non-perturbatively (DBW2 [130], [45] or Iwasaki [83]) one hopes to bring the lattice action closer to the continuum at finite lattice spacing and reduce lattice artefacts. In this work the Iwasaki gauge action is used, not in the hope that it will reduce discretization effects but rather because in combination with our fermion formulation it reduces chiral symmetry breaking effects, m_{res} , while maintaining a good rate of topological tunneling [5].

1.4 Domain Wall Fermions

The continuum QCD Lagrangian, in the massless limit, has a global chiral symmetry $\psi \rightarrow e^{i\alpha\gamma_5}\psi$ which can be neatly summarised by the vanishing of the anti-commutator of the Dirac operator with γ_5 ,

$$\{D, \gamma_5\} = 0. \quad (1.38)$$

The Wilson Dirac operator does not have this property, even at zero mass the Wilson term breaks chiral symmetry explicitly. This is related to a fundamental theorem of Neilsen and Ninomiya [111] which states, under reasonable assumptions, that an exactly chiral Dirac operator free from doublers is impossible. This fact is intimately related to the so-called QCD anomaly. This is the effect caused by the non-trivial Jacobian generated by the chiral transformation in the quantum theory, the chiral symmetry is only a symmetry of the Lagrangian and not of the path integral [67]. This leads to an anomalous divergence and non-conservation of the axial current generated by the global chiral symmetry. A chiral lattice theory would exactly conserve the axial current at finite lattice spacing and the continuum limit of such a lattice theory would fail to reproduce the correct QCD anomaly. In the continuum limit of the naive Dirac operator the doubler modes exactly cancel the anomaly [90].

A way to be free of doublers and still have almost exact chiral symmetry was first proposed by Kaplan [89] using an idea of Callan and Harvey [34] and the formulation used in practical simulations by Furman and Shamir [68] and reviewed in [86]. The idea is to consider a five dimensional theory of quarks, in an infinite volume, with a Dirac operator

$$D_5 = \not{\partial}_5 + m(s) = \gamma_\mu \partial_\mu + \gamma_5 \partial_5 + m(s) \quad (1.39)$$

where s labels the fifth coordinate and the mass is a step function in the fifth dimension

$$m(s) = \theta(s)m_0. \quad (1.40)$$

Massless states bound to the wall at $s = 0$ have to satisfy

$$D_5 \Psi = \gamma_\mu \partial_\mu \Psi = 0 \quad (1.41)$$

implying that Ψ is a combination of a four dimensional massless chiral fermion $e^{ipx}u_\pm$ with $\gamma_5 u_\pm = \pm u_\pm$ and a solution to,

$$[\pm \partial_s + m(s)]\phi_\pm(s) = 0 \quad (1.42)$$

which is solved by

$$\phi_\pm(s) = \exp\left(\mp \int_0^s m(s')ds'\right). \quad (1.43)$$

Only ϕ_+ is normalizable and thus there is a single positive chirality fermion bound to the mass defect at $s = 0$ and exponentially decaying away from it.

This is the Callan Harvey argument. Kaplan proposed a discretization of this model. With naive fermions we once again find doubler modes and do not properly describe QCD. However, by adding a $5d$ Wilson term and providing (for $p \rightarrow 0$) the mass satisfies

$$0 < m_0 < 2 \quad (1.44)$$

only one normalizable solution exists. This leaves us with an infinite lattice where there is a single chiral fermion bound to the mass defect and exponentially decaying away from it. Making the lattice finite and applying periodic boundary conditions in the fifth direction gives a second domain wall which has a fermion of opposite chirality bound to it. The two solutions overlap due to the finite fifth dimension but with exponentially suppressed amplitude.

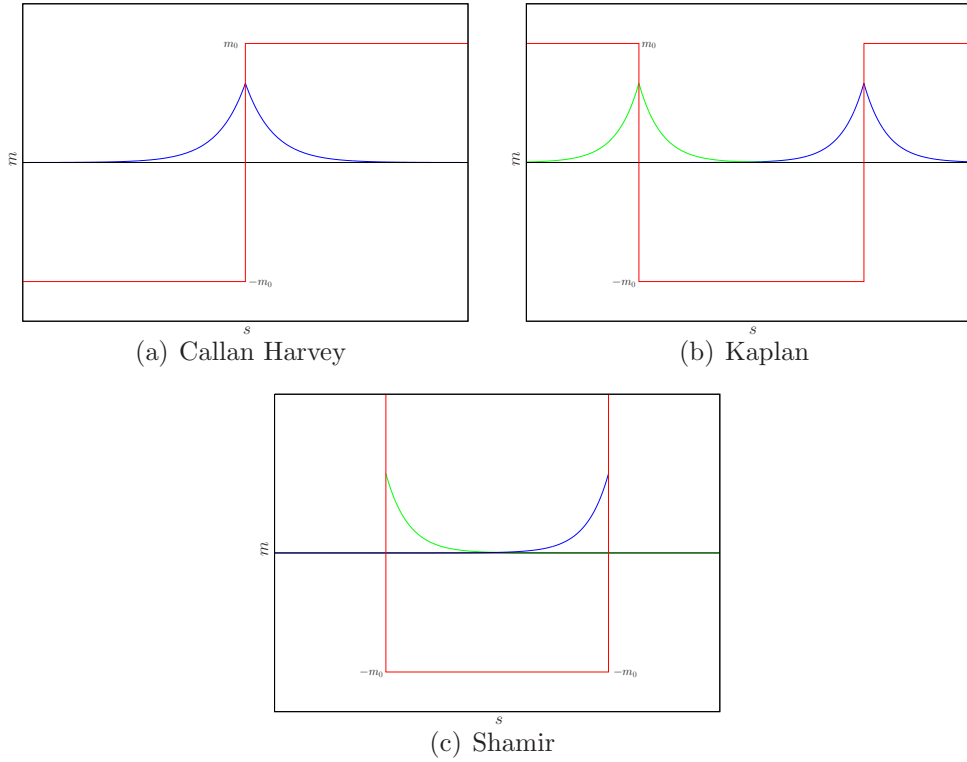


Figure 1.2: Progression of the domain wall fermion formulation from a massless state bound to a defect (a) to opposite chirality pairs bound to distant walls (b) to chiral fermions interacting only through exponentially suppressed tails (c).

Shamir [120] then proposed to cut the link connecting sites $s = 0$ and $s = L_s - 1$ (L_s is always even in this thesis), this is analagous to complete reflection at the boundaries or an infinitely high potential barrier compared to the finite barrier we had before. We will have no overlap between the modes on each wall except for the exponentially small one due to the finite size of the fifth direction. Real QCD contains massive Dirac fermions with both chirality components. If we allow the two walls to couple with strength proportional to fermion mass the two Weyl fermions will form a massive Dirac fermion. The final form of the Dirac operator is,

$$D_{DWF}(s, t) = \begin{cases} m\delta(Ls - 1, t)P_L + \delta(s, t)(D_W + 1) - \delta(s + 1, t)P_R & s = 0 \\ -\delta(s - 1, t)P_L + \delta(s, t)(D_W + 1) - \delta(s + 1, t)P_R & 0 < s < L_s - 1 \\ -\delta(s - 1, t)P_L + \delta(s, t)(D_W + 1) + m\delta(0, t)P_R & s = L_s - 1 \end{cases} \quad (1.45)$$

Here m is the mass of the Dirac fermion $P_{L,R} = \frac{1 \pm \gamma_5}{2}$ and $D_W = D_W(-M_5)$ is the Wilson operator with negative mass, equation 1.24. M_5 is the “height” of the domain wall and is tuned to give small chiral symmetry breaking, the optimal value depends on the gauge action. Writing out the matrix in $L_s \times L_s$ form looks like,

$$D_{DWF} = \begin{pmatrix} D_W + 1 & -P_L & 0 & \dots & 0 & mP_R \\ -P_R & D_W + 1 & -P_L & \dots & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & \ddots & -P_L \\ mP_L & 0 & \dots & 0 & -P_R & D_W + 1 \end{pmatrix}. \quad (1.46)$$

The four dimensional fermion fields $q(x)$ are defined from the five dimensional ones $\Psi(x, s)$ via,

$$\begin{aligned} q(x) &= P_L \Psi(x, 0) + P_R \Psi(x, L_s - 1) \\ \bar{q}(x) &= \bar{\Psi}(x, L_s - 1) P_L + \bar{\Psi}(x, 0) P_R \end{aligned} \quad (1.47)$$

and the four dimensional propagator is

$$S = P_R D^{-1}(L_s - 1, 0) P_R + P_R D^{-1}(L_s - 1, L_s - 1) P_L + P_L D^{-1}(0, 0) P_R + P_L D^{-1}(0, L_s - 1) P_L \quad (1.48)$$

Just as D_W is γ_5 Hermitian, D_{DWF} is easily shown to be $\Gamma_5 = \gamma_5 \mathcal{R}_5$ Hermitian ie.

$$\gamma_5 \mathcal{R}_5 D_{DWF} \mathcal{R}_5 \gamma_5 = D_{DWF}^\dagger, \quad (1.49)$$

the matrix $\mathcal{R}_5$ is a reflection in the fifth dimension,

$$\mathcal{R}_5(x, s; y, t) = \delta(x, y) \delta(s, L_s - 1 - t) \quad (1.50)$$

1.4.1 Chiral Symmetry

The domain wall action has a vector flavour symmetry $\psi \rightarrow e^{i\alpha^a \frac{\lambda^a}{2}} \psi$ where λ^a is a flavour rotation matrix. This symmetry gives a five dimensional conserved current

$$\begin{aligned} j_\mu(x, s) &= \frac{1}{2} (\bar{\psi}(x + \hat{\mu}, s) (1 + \gamma_\mu) U_\mu^\dagger(x + \hat{\mu}) \lambda_a \psi(x, s) - \bar{\psi}(x, s) (1 - \gamma_\mu) U_\mu(x) \lambda_a \psi(x + \hat{\mu}, s)) \\ j_5(x, s) &= \frac{1}{2} (\bar{\psi}(x, s + 1) (1 + \gamma_5) U_5^\dagger(x) \lambda_a \psi(x, s) - \bar{\psi}(x, s) (1 - \gamma_5) U_5(x) \lambda_a \psi(x, s + 1)) \end{aligned} \quad (1.51)$$

where $U_5(x) = 1$ but is written out to show that the current is exactly what one expects from a five dimensional theory. This five dimensional current satisfies a continuity equation from which it can be shown, [68], that

$$\mathcal{V}_\mu^a(x) = \sum_{s=0}^{L_s-1} j_\mu^a(x, s) \quad (1.52)$$

is conserved. Using the four dimensional quark fields of equation 1.47 we can also define the current,

$$V_\mu^a(x) = \bar{q}(x) \lambda_a \gamma_\mu q(x) \quad (1.53)$$

which is related to $\mathcal{V}$ by,

$$Z_V V_\mu = \mathcal{V}_\mu. \quad (1.54)$$

Similarly, there is an axial symmetry $\psi \rightarrow e^{i\alpha^a \frac{\lambda^a}{2} q(s)} \psi$ with

$$q(s) = \begin{cases} 1 & 0 \leq s \leq \frac{L_s}{2} - 1 \\ -1 & \frac{L_s}{2} \leq s < L_s \end{cases}$$

The corresponding axial current,

$$\mathcal{A}_\mu^a(x) = - \sum_{s=0}^{L_s-1} q(s) j_\mu^a(x, s) \quad (1.55)$$

is conserved for $m = 0$ and infinite L_s . It is related to the axial current constructed from the four dimensional fields

$$A_\mu^a(x) = \bar{q}(x) \lambda_a \gamma_\mu \gamma_5 q(x) \quad (1.56)$$

by a multiplicative renormalization,

$$Z_A A_\mu = \mathcal{A}_\mu. \quad (1.57)$$

With exact chiral symmetry $Z_A = Z_V$ [121]. $\mathcal{A}_\mu$ satisfies the continuity equation,

$$\Delta_\mu \mathcal{A}_\mu^a(x) = 2m J_5^a(x) + 2J_{5q}^a(x), \quad (1.58)$$

where Δ_μ is a backward difference operator, $J_5^a(x) = j_5^a(x, L_s)$ and $J_{5q}^a(x)$, the mid-point operator, is equal to $j_5^a(x, L_s/2 - 1)$. The midpoint term is due to chiral symmetry breaking at finite L_s . The axial Ward identities expressing chiral symmetry are therefore also altered by the mid-point term,

$$\Delta_\mu \langle \mathcal{A}_\mu^a(x) \mathcal{O}(y) \rangle = 2m \langle J_5^a(x) \mathcal{O}(y) \rangle + 2 \langle J_{5q}^a(x) \mathcal{O}(y) \rangle + i \langle \delta^a \mathcal{O}(y) \rangle. \quad (1.59)$$

In the continuum limit the usual Ward identities must be restored. The two J_5 terms must combine, changing m to give an effective quark mass,

$$m_{eff} = m + m_{res} \quad (1.60)$$

where we have put

$$J_{5q} = m_{res} J_5(x) - \frac{(Z_A - 1)}{2} \partial_\mu \mathcal{A}_\mu(x) \quad (1.61)$$

The $\partial_\mu \mathcal{A}_\mu$ term is allowed by dimensional analysis [121]. Then we recover the continuum Ward identities by renormalizing $\mathcal{A}_\mu \rightarrow Z_A \mathcal{A}_\mu$ and the quark mass $m \rightarrow m + m_{res}$. In practice $(Z_A - 1)$ has usually been argued to be small and has been neglected giving $Z_A = 1$, however recently [10] this has been found to be false at the 1% scale in current simulations and the effect of $Z_A \neq 1$ has been accounted for as an extra systematic error.

The quantity

$$m'_{res} = \frac{\langle 0 | J_{5q} | \pi \rangle}{\langle 0 | J_5 | \pi \rangle} \quad (1.62)$$

is used in practice which differs from m_{res} by $(Z_A - 1)m$ effects. The axial Ward identity gives,

$$m + m'_{res} = \frac{\langle 0 | \partial_\mu \mathcal{A}_\mu | \pi \rangle}{2 \langle 0 | J_5 | \pi \rangle} \quad (1.63)$$

sometimes called the Ward identity mass. This vanishes as the pion mass goes to zero and thus it is natural to define the chiral limit of domain wall fermions as

$$m = -m'_{res}. \quad (1.64)$$

At non-zero quark mass, ignoring the $O(Z_A - 1)$ corrections to the mid-point term, the Ward identity reads

$$Z_A \Delta_\mu \langle A_\mu^a(x) \mathcal{O}(y) \rangle = (m + m_{res}) \langle J_5^a(x) \mathcal{O}(y) \rangle + i \langle \delta^a \mathcal{O}(y) \rangle. \quad (1.65)$$

In terms of four dimensional fields,

$$J_5^a(x) = Z_P \bar{q}(x) \lambda^a \gamma_5 q(x) = Z_P P^a(x) \quad (1.66)$$

$J_5^a(x)$ is just the pseudoscalar operator $P(x)$. The scalar operator,

$$S^a(x) = \bar{q}(x) \lambda^a q(x) \quad (1.67)$$

is just like a mass term and Ward identities between the scalar vertex and the

propagator imply,

$$Z_m = \frac{1}{Z_S}. \quad (1.68)$$

Further, with exact chiral symmetry we have,

$$Z_S = Z_P. \quad (1.69)$$

Assuming that the five dimensional axial current $\mathcal{A}_\mu$ is exactly conserved, we can compute Z_A from a ratio of matrix elements of $\mathcal{A}_\mu$ and A_μ , for example,

$$\begin{aligned} C(t+1/2) &= \sum_{\vec{x}} \langle \mathcal{A}_\mu(x) P(0) \rangle \\ L(t) &= \sum_{\vec{x}} \langle A_\mu(x) P(0) \rangle \end{aligned} \quad (1.70)$$

The ratio $\frac{C(t+1/2)}{L(t)} = Z_A + O(a)$. Usually the order a errors are cancelled by considering more complicated ratios [23],

$$\frac{1}{2} \left(\frac{C(t+1/2) - C(t-1/2)}{2L(t)} + \frac{2C(t+1/2)}{L(t) + L(t+1)} \right) \quad (1.71)$$

this is the method that is used in practice.

1.5 Overlap Fermions

An alternative formulation of chiral fermions on the lattice begins by replacing equation 1.38 with

$$\{D, \gamma_5\} = 2aD\gamma_5D. \quad (1.72)$$

This is the Ginsparg Wilson relation [71] defining chiral symmetry at finite lattice spacing, the factor of the lattice spacing on the right hand side means this tends to zero in the continuum limit. For simplicity, in the remainder of this section we choose units such that a is equal to one. The chiral transformation becomes,

$$\psi \rightarrow \exp(i\alpha\gamma_5(1-D))\psi, \quad \bar{\psi} \rightarrow \bar{\psi} \exp(i\alpha(1-D)\gamma_5) \quad (1.73)$$

for a Dirac operator satisfying 1.72. The decomposition of a fermion field into left and right components is,

$$\begin{aligned}\psi_{L/R} &= \hat{P}_{L/R}\psi \\ \bar{\psi}_{L/R} &= \bar{\psi}P_{R/L}\end{aligned}\tag{1.74}$$

with

$$\hat{P}_{L,R} = \frac{1 \pm \gamma_5(1 - 2D)}{2}.\tag{1.75}$$

At finite mass we add a mass term to the action which modifies the Dirac operator:

$$\begin{aligned}m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R) &= m\bar{\psi}(P_L\hat{P}_L + P_R\hat{P}_R)\psi \\ &= m\bar{\psi}(1 - D)\psi\end{aligned}\tag{1.76}$$

$$D_m = D + m(1 - D) = m + (1 - m)D.\tag{1.77}$$

An operator satisfying the Ginsparg-Wilson relation, called the overlap operator, has been constructed [109], [110] and is given by,

$$D_O = \frac{1}{2}(1 + \gamma_5 \text{sign}(H)) = \frac{1}{2}\left(1 + \gamma_5 \frac{H}{\sqrt{H^2}}\right),\tag{1.78}$$

where the kernel of the sign function $H = \gamma_5 D$ and D is a γ_5 Hermitian Dirac operator. The simplest choice is $D = D_W(-(1 + s))$ with $|s| < 1$, the value of s affects the locality of the operator [73] and $(1 + s)$ is the analogue of the domain wall height M_5 .

1.6 Equivalence of Overlap and Domain Wall

This section outlines a series of transformations [27], [56] that show how the five dimensional formulation is equivalent to overlap fermions with a tanh approximation to the sign function and a kernel other than D_W .

First define the basis permutation matrices

$$\mathcal{P} = \begin{pmatrix} P_L & P_R & 0 & \dots & 0 \\ 0 & P_L & P_R & \dots & 0 \\ & & \dots & & \\ 0 & 0 & \dots & P_L & P_R \\ P_R & 0 & \dots & 0 & P_L \end{pmatrix} \quad (1.79)$$

with inverse $\mathcal{P}^{-1} = \mathcal{P}^\dagger$. The four dimensional propagator can be written in terms of the five dimensional DWF matrix as,

$$S(x, y) = \{\mathcal{P}^{-1} D_{DWF}(x, y; m) \mathcal{R}_5 \mathcal{P}\}_{11} \quad (1.80)$$

the indices, 11, refer to the 5th co-ordinate. Starting from the action,

$$S_{DWF} = \bar{\psi} D_{DWF} \psi \quad (1.81)$$

we go to new variables $\chi = \mathcal{P}^{-1} \psi$, $\bar{\chi} = \bar{\psi} \gamma_5 Q_L$. $Q_{L,R} = H_W P_{L,R} \pm 1$ and the action becomes,

$$S_{DWF} = \bar{\chi} D_\chi \chi \quad (1.82)$$

$$D_\chi = Q_L^{-1} \gamma_5 D_{DWF} \mathcal{P} \quad (1.83)$$

After some rearrangement, [56], we find

$$D_\chi = \begin{pmatrix} P_L - mP_R & -T^{-1} & 0 & \dots & 0 \\ 0 & 1 & -T^{-1} & \dots & \vdots \\ \vdots & & \dots & & \\ 0 & 0 & \dots & 1 & -T^{-1} \\ -T^{-1}(P_R - mP_L) & 0 & \dots & 0 & 1 \end{pmatrix} \quad (1.84)$$

$$T^{-1} = -Q_L^{-1} Q_R = \frac{1 + H_T}{1 - H_T} \quad (1.85)$$

$$H_T = \frac{1}{2 + H_W \gamma_5} H_W \propto H_W + O(a_5), \quad (1.86)$$

a_5 is the lattice spacing in the fifth dimension and the Hermitian operator H_T is called the Shamir Kernel. This matrix can then be factored around its top left

element by taking a Schur compliment,

$$M = \begin{pmatrix} D & C \\ B & A \end{pmatrix} = \begin{pmatrix} 1 & CA^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & A \end{pmatrix} \begin{pmatrix} 1 & 0 \\ A^{-1}B & 1 \end{pmatrix} \quad (1.87)$$

$$S = D - CA^{-1}B \quad (1.88)$$

$$\det(M) = \det(A) \det(S) \quad (1.89)$$

For our matrix,

$$D = P_L - mP_R \quad (1.90)$$

$$A = \begin{pmatrix} 1 & -T^{-1} & 0 & \dots & \vdots \\ 0 & 1 & -T^{-1} & \dots & \vdots \\ \vdots & \vdots & \ddots & \dots & \vdots \\ 0 & 0 & \dots & 1 & -T^{-1} \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} \quad (1.91)$$

$$A^{-1} = \begin{pmatrix} 1 & T^{-1} & T^{-2} & \dots & T^{-L_s+2} \\ 0 & 1 & T^{-1} & \dots & T^{-L_s+3} \\ \vdots & \vdots & \ddots & \dots & \vdots \\ 0 & 0 & \dots & 1 & T^{-1} \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} \quad (1.92)$$

$$C = (T^{-1}, 0, \dots, 0) \quad (1.93)$$

$$B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -T^{-1}(P_R - mP_L) \end{pmatrix} \quad (1.94)$$

$$S = -[T^{-L_s} + 1]\gamma_5 \times \left[\frac{1+m}{2} + \frac{1-m}{2} \gamma_5 \frac{T^{-L_s} - 1}{T^{-L_s} + 1} \right] \quad (1.95)$$

We will see in section 1.7 how to cancel the factor $[T^{-L_s} + 1]\gamma_5$. The other term contains an approximation to the sign function,

$$\frac{T^{-L_s} - 1}{T^{-L_s} + 1} = \tanh \left(-\frac{L_s}{2} \log |T| \right) \quad (1.96)$$

Perhaps more transparently this is,

$$\frac{T^{-L_s} - 1}{T^{-L_s} + 1} = \frac{(1 + H_T)^{L_s} - (1 - H_T)^{L_s}}{(1 + H_T)^{L_s} + (1 - H_T)^{L_s}} = \epsilon_{L_s/2}(H_T) \quad (1.97)$$

using Neuberger's polar decomposition approximation to the sign function [108].

Thus the operator

$$D_{ov} = \left[\frac{1 + m}{2} + \frac{1 - m}{2} \gamma_5 \frac{T^{-L_s} - 1}{T^{-L_s} + 1} \right] \quad (1.98)$$

looks very much like the massive overlap operator, equation 1.77. Figure 1.3 shows the spectra of the exact and effective overlap operators in the unit gauge for two choices of kernel: the Shamir kernel, which is used exclusively in this thesis; and the Wilson kernel, typical of 4d approaches to overlap fermions.

Let us establish the connection to the usual DWF propagator defined in terms of ψ fields.

$$\begin{aligned} D_{ov} &= [D_\chi^{-1}(m=1)D_\chi(m)]_{11} \\ &= [\mathcal{P}^{-1}D_{DWF}^{-1}(m=1)D_{DWF}(m)\mathcal{P}]_{11} \end{aligned} \quad (1.99)$$

A contact term, [58], is subtracted from the inverse to make the physical propagator,

$$\begin{aligned} \tilde{D}_{ov}^{-1} &= \frac{1}{1 - m} [D_{ov}^{-1} - 1] \\ &= \frac{1}{1 - m} [\mathcal{P}^{-1}D_{DWF}^{-1}(m)D_{DWF}(1)\mathcal{P} - 1]_{11} \\ &= \frac{1}{1 - m} [\mathcal{P}^{-1}D_{DWF}^{-1}(m)D_{DWF}(1)\mathcal{P} - \mathcal{P}^{-1}D_{DWF}^{-1}(m)D_{DWF}(m)\mathcal{P}]_{11} \\ &= \frac{1}{1 - m} [\mathcal{P}^{-1}D_{DWF}^{-1}(m)[D_{DWF}(1) - D_{DWF}(m)]\mathcal{P}]_{11} \end{aligned} \quad (1.100)$$

Using,

$$[D_{DWF}(1) - D_{DWF}(m)]_{s,t} \mathcal{P}_{t,1} = \quad (1.101)$$

$$(1-m) \begin{pmatrix} 0 & 0 & \dots & 0 & P_R \\ 0 & 0 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 0 & 0 \\ P_L & 0 & \dots & 0 & 0 \end{pmatrix} \begin{pmatrix} P_L \\ 0 \\ \vdots \\ 0 \\ P_R \end{pmatrix} = (1-m) \begin{pmatrix} P_R \\ 0 \\ \vdots \\ 0 \\ P_L \end{pmatrix} = (1-m) \mathcal{R}_5(s,t) \mathcal{P}(t,1)$$

gives

$$\tilde{D}_{ov}^{-1} = [\mathcal{P}^{-1} D_{DWF}^{-1}(m) \mathcal{R}_5 \mathcal{P}]_{11} \quad (1.102)$$

which is just the same as 1.48, the usual domain wall propagator. By modifying the five dimensional formulation different kernels and different approximations to the sign function may be introduced [91]. In this sense the physical picture described in section 1.4 is gone and the $5d$ matrices just appear as convenient ways to describe different types of overlap operator. In practice working with the $5d$ formulation is faster since evaluating the sign function is difficult and we can exploit the sparsity of D_{DWF} for linear solvers.

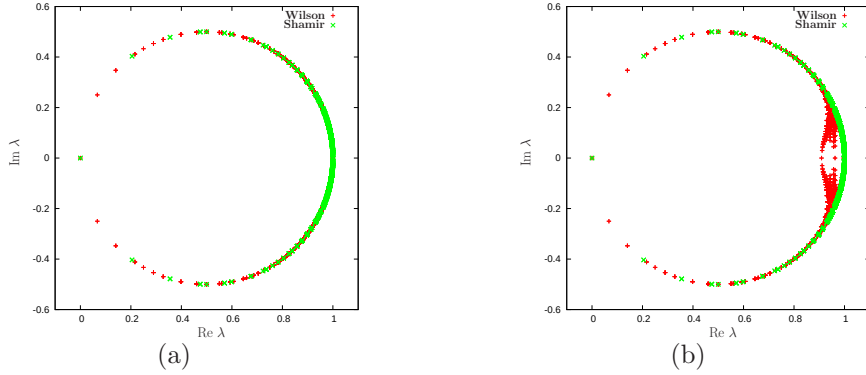


Figure 1.3: (a) The spectrum of the free overlap operator 1.78 using the Wilson Kernel or the Shamir Kernel with exact evaluation of the sign function. (b) The $4d$ effective overlap operator 1.99 derived from $L_s = 8$ domain wall fermions with both kernels. Comparing overlap to domain wall, the low modes of each kernel match rather closely, the difference is mainly in the high part of the spectrum, especially for the Wilson kernel which is a worse approximation for smaller L_s , at least in the free field. Doubling L_s , not shown, produces a spectrum indistinguishable on this scale from the exact overlap.

1.7 Pauli-Villars Fields

The formulation of domain wall fermions we have used thus far still has a problem if the $L_s \rightarrow \infty$ limit is taken before the continuum limit, as will always be the case. If domain wall fermions are interpreted as two light chiral modes bound to the mass defects together with $O(L_s)$ heavy Wilson modes propagating through the five dimensional volume taking L_s to infinity is evidently going to be tricky [60]. In the overlap formalism this shows up in the $[T^{-L_s} + 1]\gamma_5$ factor in equation 1.95 which diverges as $L_s \rightarrow \infty$. For example, if v an eigenvector of H_T with positive eigenvalue λ , $v^\dagger T^{-1}v = \frac{1+\lambda}{1-\lambda}$, the ratio is greater than one and hence diverges when raised to higher and higher powers.

The way to cancel this bulk infinity was pointed out by [133]. Introducing new pseudo-fermion fields ϕ_{PV} , called Pauli-Villars fields, into the action, the partition function is

$$Z = \int \mathcal{D}U \mathcal{D}\phi \mathcal{D}\phi^* \det[D(m)] \exp(-S_G - S_{PV}) \quad (1.103)$$

$$S_{PV} = \phi^* D(1) \phi. \quad (1.104)$$

$D_{PV} = D_{DWF}(1)$ is just the domain wall Dirac operator with unit mass. The c-number fields ϕ can be integrated out and give,

$$Z = \int \mathcal{D}U \frac{\det[D(m)]}{\det D[1]} \exp(-S_G). \quad (1.105)$$

From 1.95 with $m = 1$ the Pauli-Villars determinant is just $[T^{-L_s} + 1]\gamma_5$ which, cancels the same factor from the fermion determinant.

Later we will wish to evaluate the effective overlap operator 1.99 often, as part of a Lanczos process, thus a way to quickly perform the inverse of the Pauli Villars operator is useful [84]. Start with the following matrix, which does a discrete Fourier transformation with anti-periodic boundary conditions in the fifth direction and does not affect the other directions,

$$U_{n,m} = \exp\left(i \frac{2\pi}{L_s} \left(m + \frac{1}{2}\right)n\right). \quad (1.106)$$

$$D_{n,m}(1)U_{m,p} = (D_W + 1 - P_L e^{i\alpha_n} - P_R e^{-i\alpha_n})U_{n,p}, \quad (1.107)$$

ie. the columns of U are eigenvectors of $D(1)$ with $\alpha_n = \frac{2\pi}{L_s}(n + \frac{1}{2})$. U is unitary so this implies,

$$\{U^\dagger D U\}_{n,m} = (D_W + 1 - P_L e^{i\alpha_n} - P_R e^{-i\alpha_n}) \delta_{n,p}. \quad (1.108)$$

The $L_s \times L_s$ diagonal matrix on the right can be written as

$$D_{TW}(n) \delta_{n,p} \quad (1.109)$$

$$D_{TW}(n) = D_W(1 - M_5 - \cos(\alpha_n)) + i\gamma_5 \sin(\alpha_n). \quad (1.110)$$

and solutions to the equation $D(1)\psi = \eta$ are,

$$\psi_n = U_{n,r}^\dagger D_{TW}^{-1}(r) U_{r,m} \eta_m. \quad (1.111)$$

We have traded the inversion of an $L_s \times L_s$ system for L_s separate inverses. This gives a modest speed up (~ 2) when inverting the Pauli-Villars operator, which is helpful for repeatedly evaluating the effective overlap, see chapter 5.

Chapter 2

Renormalization

Formulating the Standard Model, or any quantum field theory, involves constructing the Lagrangian from products of fields and derivatives of fields while satisfying symmetry principles, such as a local gauge invariance. In general there are an infinite number of terms invariant under these symmetries so an infinite number of terms might be included. In fact QCD and the Standard Model have only a finite, and in fact small, number of terms in their Lagrangians. In this chapter we discuss what happened to the other terms which will lead us to asymptotic freedom, how to take a continuum limit and how to renormalize quantities calculated in lattice simulations.

2.1 Wilson's Renormalization Group

Wilson's idea [137] is to consider the space of all possible theories satisfying certain symmetries. Following most introductory presentations, references [82], [48] for example, we consider scalar field theory in d dimensions where the Lagrangian is invariant under change of sign of the fields, $\phi \rightarrow -\phi$. First the Euclidean partition function

$$\mathcal{Z} = \int D\phi e^{-S[\phi]} \tag{2.1}$$

is made finite and well defined by introducing a regulator. Useful regulators are: a momentum cutoff, integrating over only low energy modes of ϕ ; dimensional regularization, using fewer than 4 spacetime dimensions; or a lattice. The regulator defines the path integral rigorously and calculations may be carried

out in perturbation theory or non-perturbatively. For ϕ^4 theory we will use a momentum cutoff Λ , this may be introduced by hand as an upper limit in the perturbative integrals over all momenta or it may reflect an underlying lattice formulation, in which case it is the reciprocal of the lattice spacing, $\Lambda = a^{-1}$.

The most general action satisfying the symmetry,

$$S[\phi] = S[-\phi] \quad (2.2)$$

is

$$S = \int d^d x \frac{1}{2} (\partial_\mu \phi)^2 + V(\phi) \quad (2.3)$$

$$V(\phi) = \sum_n \Lambda^{d-n(d-2)} \frac{g_{2n}}{(2n)!} \phi^{2n} + \sum_i \Lambda^{d-d_i} g_{\mathcal{O}_i} \mathcal{O}_i \quad (2.4)$$

The couplings, g , are all dimensionless and an explicit factor Λ has been included to get the dimension of each term right, the operators $\mathcal{O}_i$ contain derivatives of ϕ . An operator with m derivatives and $2n$ factors of ϕ has to be scaled by $\Lambda^{d-m-n(d-2)}$ so $d_i = m + n(d-2)$. The coupling of the kinetic term has been absorbed into the definition of ϕ , ie. the term $\frac{Z}{2}(\partial_\mu \phi)^2$ is brought to canonical form by the rescaling $\phi \rightarrow Z^{1/2} \phi$. The terms with smallest mass dimension are,

$$\frac{\Lambda^2 g_2}{2} \phi^2 + \frac{\Lambda^{(4-d)} g_4}{4!} \phi^4 \quad (2.5)$$

$\Lambda^2 g_2 = m^2$ and $\Lambda^{(4-d)} g_4 = \lambda$ in the usual notation for ϕ^4 theory.

Take the set of all the couplings $\mathcal{K}_i = \{g_2, g_4, \dots\}$. Consider a transformation $\mathcal{R}(s)$ that integrates out high modes of ϕ between $\frac{\Lambda}{s}$ and Λ and does a simple rescaling of the fields and their arguments. $\mathcal{R}(s)$ is called a renormalization group transformation (RGT). Since we have considered the most general action possible the only effect a RGT can have on the action is to change the couplings,

$$\mathcal{R}(t, \mathcal{K}_i) = \mathcal{K}'_i \quad (2.6)$$

so RGTs just transform (infinite dimensional) vectors into vectors.

Let $\mathcal{P}(g_i(\Lambda), p)_\Lambda$ be a physical quantity, like a correlation function, calculated using the theory with cutoff Λ . It depends on the couplings and some physical scales p , like particle momenta or masses. The renormalization group idea is that we can change the cutoff Λ to $\Lambda' < \Lambda$ and keep the physics at energies less than

Λ' fixed. A calculation of $\mathcal{P}$ with either cutoff must give the same result,

$$\mathcal{P}(g_i(\Lambda), p)_\Lambda = \mathcal{P}(g_i(\Lambda'), p)_{\Lambda'} \quad (2.7)$$

which is possible if the couplings change in just the right way, that is, via (exact) renormalization group transformations.

The physical process that gives $\mathcal{P}$ has typical energy scale p ; we can follow equation 2.7 down as far as possible taking the cutoff $\Lambda \rightarrow p$. Thus the running of the couplings with cutoff can be modified to give the running of the couplings with the typical energy scale of the problem. If we consider infinitesimal changes in the cutoff $\Lambda' = \Lambda + d\Lambda$ equation 2.7 becomes,

$$\Lambda \frac{d}{d\Lambda} \mathcal{P}(g_i(\Lambda), p)_\Lambda = \left(\Lambda \frac{\partial}{\partial \Lambda} + \beta_i \frac{\partial}{\partial g_i} \right) \mathcal{P}(g_i(\Lambda), p)_\Lambda = 0 \quad (2.8)$$

$$\beta_i = \Lambda \frac{\partial g_i}{\partial \Lambda}. \quad (2.9)$$

These are the beta functions for each coupling and β_{g_4} is called “the” beta function of ϕ^4 theory. The beta functions describe the how the couplings change with the cutoff in order to keep the low energy physics fixed. The beta function of the coupling g_i will be of the form,

$$\Lambda \frac{dg_i}{d\Lambda} = \tilde{\beta} + (d_i - d)g_i \quad (2.10)$$

where $\tilde{\beta}$ is a quantum correction. We ignore quantum corrections for now and consider a change, $\Lambda \rightarrow \frac{\Lambda}{s}$ with $s > 1$. This flows into the IR and along this flow the mass relative to the cutoff diverges. Thus massive particles become infinitely heavy in the IR limit and do not propagate so only theories with massless particles can have non-trivial IR limits.

There is an intuitive, geometric setting for renormalization group ideas in the space of all theories, where the coordinates are just the coupling constants. A renormalization group transformation, $\mathcal{R}$, takes a theory (ie. a set of couplings $\mathcal{K}_i$) and moves it along some trajectory in the space. The subspace of theories with massless particles and non-trivial IR limits is called the critical surface. Theories starting on the critical surface flow under RGTs and usually approach

a fixed point, that is a theory, $\mathcal{K}_i^*$, such that,

$$\mathcal{R}(s, \mathcal{K}_i^*) = \mathcal{K}_i^*. \quad (2.11)$$

One could start with a theory defined with cutoff Λ and perform RGTs with scale factor s to obtain $\mathcal{K}_i(s)$. Consider two infinitesimally separated theories,

$$\mathcal{K}_i(s(1 + \epsilon)) - \mathcal{K}_i(s) = \mathcal{R}(1 + \epsilon, \mathcal{K}_i(s)) - \mathcal{R}(1, \mathcal{K}_i(s)) \quad (2.12)$$

and Taylor expand to get,

$$\beta_i(\mathcal{K}) = s \frac{\partial \mathcal{K}_i}{\partial s} = \left. \frac{\partial \mathcal{R}(t, \mathcal{K}_i)}{\partial t} \right|_{t=1}. \quad (2.13)$$

This is equivalent to our previous definition of the beta functions, equation 2.9, after restoring a factor Λ on top and bottom.

Consider a theory near a fixed point $\mathcal{K}_i^*$

$$\mathcal{K}_i(s) = \mathcal{K}_i^* + \Delta \mathcal{K}_i(s), \quad (2.14)$$

the difference between the corresponding beta functions is,

$$\begin{aligned} s \frac{\partial \mathcal{K}_i}{\partial s} - s \frac{\partial \mathcal{K}_i^*}{\partial s} &= s \frac{\partial \Delta \mathcal{K}_i}{\partial s} \\ &= \left. \frac{d\beta_i}{d\mathcal{K}_j} \right|_{\mathcal{K}_i^*} \Delta \mathcal{K}_j \end{aligned} \quad (2.15)$$

where we have Taylor expanded the beta function and used the fact that it vanishes at the fixed point to get to the second line. This shows that,

$$s \frac{\partial \Delta \mathcal{K}_i}{\partial s} = T_{ij} \Delta \mathcal{K}_j \quad (2.16)$$

$$T_{ij} = \left. \frac{d\beta_i}{d\mathcal{K}_j} \right|_{\mathcal{K}_i^*} \quad (2.17)$$

ie. the renormalization group flow is linear in the neighbourhood of a fixed point.

We assume T_{ij} is diagonalizable and change the basis in theory space,

$$\begin{aligned}\sigma &= V \Delta \mathcal{K}_i \\ D &= V T V^{-1}\end{aligned}\tag{2.18}$$

so that D is diagonal. Then equation 2.16 becomes

$$s \frac{\partial \sigma_i}{\partial s} = D_{ii} \sigma_i = \eta_i \sigma_i \tag{2.19}$$

where η_i are the eigenvalues of T . Restoring the cutoff Λ on the left and solving gives,

$$\sigma_i(\Lambda) = \left(\frac{\Lambda}{\Lambda'} \right)^{\eta_i} \sigma_i(\Lambda'). \tag{2.20}$$

The eigenvalue η separates into a classical piece that gives the canonical scaling behaviour of the coupling and a quantum correction, ie.

$$\eta_i = \gamma_i + (d_i - d) \tag{2.21}$$

compare this to equation 2.10, which is the same thing in a different basis.

From this analysis we can see if the i^{th} coupling has:

- $\eta_i > 1$: this implies deviation from the fixed point increases under RGTs (a relevant coupling)
- $\eta_i < 1$: implies convergence towards the fixed point (irrelevant coupling)
- $\eta_i = 1$: (marginal coupling).

Finding a marginal coupling usually means we have calculated the eigenvalue in some approximation and we need to improve the approximation to discover if the coupling is relevant or irrelevant. This is the case for g_4 in ϕ^4 theory to leading order. Truly marginal directions indicate a line of fixed points.

2.1.1 ϕ^4 theory

We can perform Wilson's RGT for ϕ^4 theory, schematically at least. Start at the point in theory space where all the higher order couplings are zero and the cutoff

is Λ . The action is,

$$S = \int d^d x \frac{1}{2} (\partial_\mu \phi)^2 + \frac{m}{2} \phi^2 + \frac{\lambda}{4!} \phi^4 \quad (2.22)$$

using the usual conventions for ϕ^4 theory and letting the scale be reabsorbed by the couplings. Next, split $\phi(k) \rightarrow \phi(k) + \hat{\phi}(k)$ where $\phi(k)$ is non-zero only for $0 < k < \frac{\Lambda}{s}$ and $\hat{\phi}(k)$ is non-zero only for $\frac{\Lambda}{s} < k < \Lambda$. The action splits into the form of equation 2.22 plus

$$\begin{aligned} & \int d^d x (\partial_\mu \hat{\phi})^2 + \frac{m}{2} \hat{\phi}^2 + \frac{\lambda}{4!} \hat{\phi}^4 \\ & + \lambda \left(\frac{1}{6} \phi^3 \hat{\phi} + \frac{1}{4} \phi^2 \hat{\phi}^2 + \frac{1}{6} \phi \hat{\phi}^3 \right) \end{aligned} \quad (2.23)$$

where we used the orthogonality of ϕ and $\hat{\phi}$ remove products $\phi \hat{\phi}$. The integration measure in the partition function also changes $D\phi \rightarrow D\phi D\hat{\phi}$. When integrated the $D\hat{\phi}$ path integral generates terms proportional to every product of fields allowed by the $\phi \rightarrow -\phi$ symmetry. This gives an action,

$$\begin{aligned} S = \int d^d x & \frac{1}{2} (1 + \delta Z) (\partial_\mu \phi)^2 + \frac{1}{2} (m^2 + \delta m^2) \phi^2 \\ & + \frac{1}{4!} (\lambda + \delta \lambda) \phi^4 + \Delta C \phi^6 + \dots \end{aligned} \quad (2.24)$$

defined with the cutoff $\frac{\Lambda}{s}$. After making the change of variables $k \rightarrow k' = sk$ the new variable is integrated from 0 to Λ . The change of variables gives an overall factor s^d as well as a factor s^{-1} for every derivative. We have no coupling constant in front of the kinetic term, choosing instead to rescale the field. The kinetic term has become,

$$s^d \frac{1}{2} (1 + \delta Z) s^{-2} (\partial'_\mu \phi)^2 \quad (2.25)$$

so a rescaling of the field,

$$\phi' \rightarrow (s^{d-2} (1 + \delta Z))^{1/2} \phi \quad (2.26)$$

returns it to canonical form. The action is now,

$$S = \int d^d x' \frac{1}{2} (\partial'_\mu \phi')^2 + \frac{1}{2} (m'^2) \phi'^2 + \frac{1}{4!} \lambda' \phi'^4 + C' \phi'^6 + \dots \quad (2.27)$$

The integration and rescaling $k \rightarrow k' = sk$ is a renormalization group transformation. The form of the action is fixed but the coupling in front of each term has changed. In theory space we have moved from $\mathcal{K}_i$ to $\mathcal{K}'_i$,

$$\mathcal{K}_i = \{m^2, \lambda, 0, \dots\} \quad (2.28)$$

$$\begin{aligned} \mathcal{K}'_i &= \{m'^2, \lambda', C', \dots\} \\ &= \{(m^2 + \delta m^2)(1 + \delta Z)^{-1} s^2, \\ &\quad (\lambda + \delta \lambda)(1 + \delta Z)^{-2} s^{4-d}, \\ &\quad (C + \delta C)(1 + \delta Z)^{-3} s^{6-2d}, \dots\} \end{aligned} \quad (2.29)$$

where originally $C = 0$. A calculation with the original action 2.22 and cutoff Λ and one with the action 2.27 will give the same result because, if the integration is exact, they are the same action.

If we started with $m = 0$ (on the critical surface) and $\lambda = 0$ our transformation will take us to exactly the same point in the space of couplings because the kinetic term splits in two, the high mode part gives a multiplicative constant which can be absorbed by rescaling the fields. All the couplings initially at zero stay there. This called a Gaussian fixed point. QCD too has a Gaussian fixed point at zero quark mass and zero gauge coupling. Near the fixed point the RGT is approximately linear which means that for example $g'_2 = c_1 g_2 + c_2 g_4 + c_3 g_6 + \dots$ in this region. The simplest approximation is to ignore quantum corrections (the δ terms) then, in four dimensions,

$$\begin{aligned} m'^2 &= m^2 s^2 \\ \lambda' &= \lambda \\ C' &= s^{-2} C \text{ etc.} \end{aligned} \quad (2.30)$$

the scaling is simply given by the canonical scaling and we see the mass is a relevant coupling, λ is marginal and the rest are irrelevant and unless the quantum

corrections become very large this classification holds. Including quantum corrections would show us that λ was marginally irrelevant and would modify the running of mass slightly for small RGTs [114]. We can also calculate the beta functions by considering an infinitesimal RGT, see [82], where we find, including lowest order quantum corrections,

$$\beta(\lambda) = \frac{3\lambda^2}{16\pi}. \quad (2.31)$$

The important thing for us is that the sign is positive, indicating that the coupling is large for large scales and small for small scales. We now turn to QCD which is our primary concern.

2.2 QCD and the continuum limit

The renormalization group acting on QCD has to respect gauge invariance. This means the structure of the covariant derivative must be preserved. We rescale the gauge fields as in chapter 1 changing the covariant derivative and the gauge kinetic term,

$$A_\mu \rightarrow \frac{A_\mu}{g} \quad (2.32)$$

$$D_\mu = \partial_\mu + igA_\mu \rightarrow \partial_\mu + iA_\mu \quad (2.33)$$

$$S_G[A] \rightarrow \frac{1}{4g^2} F_{\mu\nu} F_{\mu\nu} \quad (2.34)$$

with the usual definition of the field strength $F_{\mu\nu}$. Now under RGTs the renormalization of the gauge coupling is just a further renormalization of the kinetic term and the covariant derivative is preserved. Regularizations not respecting gauge invariance are ruled out, RGTs would generate gauge non-invariant terms changing the form of the action. A simple momentum cutoff as we had before is not possible, but a lattice formulation or a dimensional regularization is allowed.

Politzer, Gross and Wilczek, references [115], [79], calculated the beta function of QCD. It turns out to be, to leading order with n_f fermions,

$$\beta = -\frac{g^3}{(4\pi)^2} \left(11 - \frac{2}{3}n_f \right) = -\beta_0 g^3. \quad (2.35)$$

For the situation of interest to us $n_f \leq 6$ and $\beta_0 > 0$. In order to keep with the usual notation for QCD μ is the typical scale of interactions, see section 2.3. The renormalization group equation,

$$\beta = \mu \frac{\partial g}{\partial \mu} = -\beta_0 g^3. \quad (2.36)$$

is solved by,

$$g^2(\mu) = \frac{1}{c + 2\beta_0 \log(\mu)} = \frac{1}{2\beta_0 \log(\mu/\Lambda_{QCD})}. \quad (2.37)$$

The integration constant c has been rewritten as a dimensionful parameter Λ_{QCD} (nothing to do with the Λ of the previous section) that depends on the renormalization conditions. For QCD $\beta_0 > 0$, so in the UV regime, $\mu \rightarrow \infty$, $g(\mu) \rightarrow 0$. This property is called asymptotic freedom. The other possibility, realised in ϕ^4 theory, QED or in QCD with large n_f , is $\beta_0 < 0$. In this case $g(\mu) \rightarrow \infty$ as the cutoff is taken to infinity. This tells us that, at a certain scale $\sim \Lambda_{QCD}$, the coupling becomes too large to use perturbation theory. This is also the case for QCD in the IR, $\mu \rightarrow 0$ limit. Hence the need for non-perturbative calculations of low energy physics.

The quark mass also flows under RGTs. Writing the renormalization group equation for the quark mass,

$$\mu \frac{dm}{d\mu} = m\gamma(g(\mu)) = -m(\gamma_0 g^2 + \dots) \quad (2.38)$$

the function γ is called the anomalous dimension and gives the modification by quantum corrections of the scaling behaviour of the mass. To lowest order in g the solution is,

$$m = m_0 g^{\frac{\gamma_0}{\beta_0}}, \quad (2.39)$$

as the coupling flows to zero, so does the quark mass ($\gamma_0 > 0$) [42].

The masses and strong coupling constant are the only relevant couplings for QCD. There is a fixed point at $g = 0$, $m = 0$. We show the flows around the fixed point in a subspace with one irrelevant direction (there are actually an infinite number) and one relevant direction, g , in figure, 2.1. The flow coming out of the fixed point is called the renormalized trajectory. All other flows converge on this trajectory deep in the IR. This is the notion of universality; there are only a small number of relevant couplings (for massless QCD only g is relevant) and an infinite

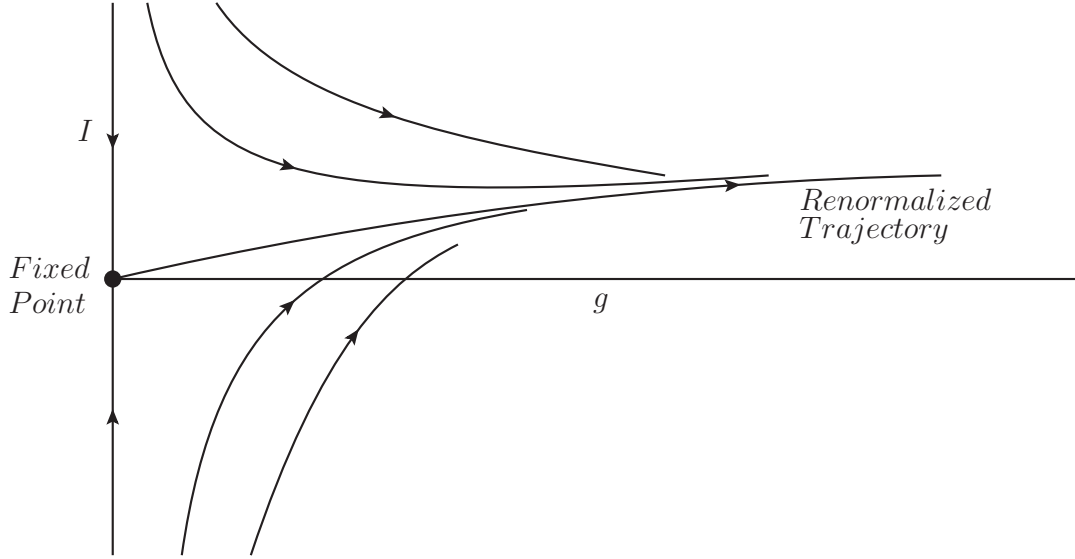


Figure 2.1: Flows with one relevant (g) and one irrelevant (I) direction showing the fixed point and the renormalized trajectory. The arrows point in the direction of RG flow which is into the IR.

number of irrelevant ones. Thus the domain of attraction of the fixed point is infinite dimensional and theories in the IR are determined by specifying only one parameter, which tells us where we are along the renormalized trajectory. Only theories on the renormalized trajectory are defined on all length scales, figure 2.1. At high scales (left of the graph) is a theory of almost free quarks and gluons, at low scales (right of the graph) is a complicated low energy theory (of meson and baryon fields for example). Off the renormalized trajectory theories flow to the same low energy effective theory but at high scales the irrelevant couplings diverge and the cutoff cannot be taken to infinity while keeping the theory finite - these are non-renormalizable theories.

In a lattice regularization the continuum limit is the process of taking the lattice spacing finer and finer, $a \rightarrow 0$ or the cutoff $a^{-1} \rightarrow \infty$, while keeping the low energy physics fixed. To lie exactly on the renormalized trajectory would require knowing an infinite set of irrelevant couplings. In practice what we do is simulate a lattice theory with some values of the coupling and some unknown irrelevant parameters due to discretization errors, points (a) and (b) in figure 2.2. Figure 2.2 shows the case where the quark masses have been taken to zero, the

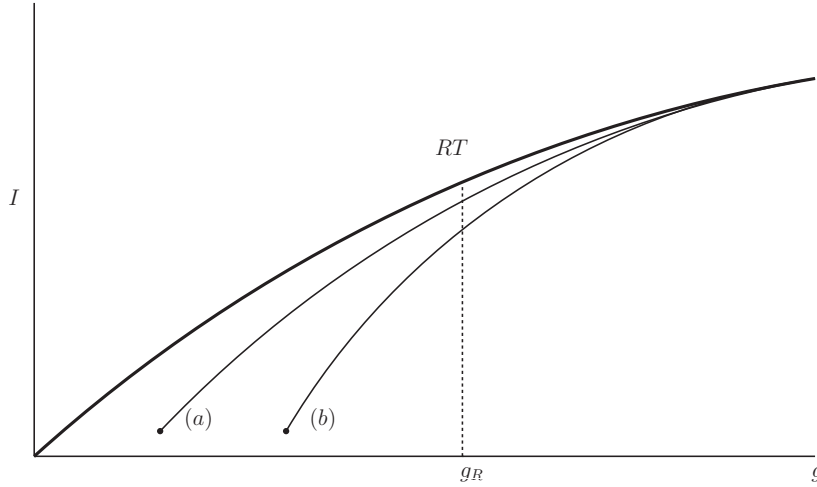


Figure 2.2: The renormalized trajectory (RT) and the trajectories that would be obtained starting with a given lattice action and doing RGTs (a) and (b). They converge to the RT at low energy due to universality. At low energy where $g = g_R$ deviation from the RT is small and results can be reliably extrapolated.

chiral limit of QCD. The renormalized trajectory here is not the physical one but an approximation where the pion mass vanishes. In fact for massless QCD the one relevant parameter is absorbed into the renormalization group dependence and we just need to fix one dimensionful quantity to set the units, then all other quantities are determined. In the simulated lattice theories we calculate observables at scales much lower than the cutoff, so we are close to the renormalized trajectory, and extrapolate to the limit of vanishing initial coupling, giving us the value we would get if we had used the parameters on the renormalized trajectory. This is much easier than trying to simulate on the renormalized trajectory with a “perfect action” though attempts have been made, [20].

If non-zero quark masses are to be included, as in [10] for example, which renormalized trajectory we converge to can be specified by the fixing the values of a number of physical quantities equal to the number of relevant parameters; for example that the ratio of the proton mass to the omega mass should equal its experimentally measured value. The quark masses have to be tuned so that the trajectory starting from the initial bare coupling lies close to the correct renormalized trajectory and so that the physical ratios are equal to the experimentally measured ones. This tuning is carried out for a number of bare couplings (ie. points (a), (b), ...) and the continuum limit takes us onto the

renormalized trajectory that reproduces the correct physics, with the ratios such as $\frac{m_p}{m_\Omega}$ equal to their experimentally determined values.

2.3 Callan Symanzik Equation and Scaling

We will now relate the Wilsonian picture to conventional, perturbative renormalization. When a theory is perturbatively renormalized one starts with the simplest action and, as infinities are encountered in perturbation theory, counter terms are added to cancel these. The intricacies of perturbation theory for renormalizable field theories mean that all the divergences can be cancelled by counter terms of the same form as the couplings in the original, simple action. The action plus counter terms is called the renormalized action, written in terms of renormalized fields and is equal to the bare action, written in terms of bare fields, in a similar way that theories with different cutoffs were related in section 2.1.1.

In fact for ϕ^4 theory we have,

$$S = \int d^d x \frac{1}{2} (\partial_\mu \phi_R)^2 + \frac{m_R^2}{2} \phi_R^2 + \frac{\lambda_R}{4!} \phi_R^4 \quad (2.40)$$

$$S_{ct} = \int d^d x \frac{\delta Z}{2} (\partial_\mu \phi_R)^2 + \frac{\delta m^2}{2} \phi_R^2 + \frac{\delta \lambda}{4!} \phi_R^4 \quad (2.41)$$

as we found on doing an RGT explicitly.

$$S_B[\phi_B] = S_R[\phi_R] = S[\phi_R] + S_{ct}[\phi_R] \quad (2.42)$$

implying,

$$m_B^2 = m_R^2 + \delta m^2 \quad (2.43)$$

$$\lambda_B = \lambda_R + \delta \lambda$$

and the renormalized field is related to the bare field by,

$$\phi_B = Z^{1/2} \phi_R \quad (2.44)$$

$$Z = (1 + \delta Z).$$

The infinities are removed by means of renormalization conditions. These are

relations that we enforce on the vertex functions of the theory at some physical scale μ , we give many specific schemes later in section 4.1.1. The appearance of a scale μ means the vertex functions are μ dependent.

The n -point vertex functions $\Gamma^{(n)}(p_i; g, \mu)$ contain n fields so,

$$\Gamma_B^{(n)}(p_i; g_B) = Z^{-n/2}(\mu) \Gamma_R^{(n)}(p_i; g_R, \mu) \quad (2.45)$$

(the sign of the exponent of Z comes from Fourier transforming). The bare action doesn't know about μ , so the combination of Z and $\Gamma^{(n)}$ must conspire to be μ independent. This is expressed by the Callan-Symanzik equation,

$$\begin{aligned} \mu \frac{d}{d\mu} \Gamma_B^{(n)} &= 0 \\ \Rightarrow \left(\mu \frac{\partial}{\partial \mu} + \beta \frac{\partial}{\partial g_R} - n\gamma \right) \Gamma_R^{(n)} &= 0 \end{aligned} \quad (2.46)$$

$$\beta = \mu \frac{\partial g_R}{\partial \mu} \quad (2.47)$$

$$\gamma = \frac{1}{2} \mu \frac{\partial \ln Z}{\partial \mu}. \quad (2.48)$$

These are not the same β 's as before. Previously we looked at changing the cutoff not the renormalization scale, so we had $\frac{\partial}{\partial a}$ instead of $\frac{\partial}{\partial \mu}$. The two are related in perturbation theory, the first two coefficients of the expansion in g are the same because the bare and renormalized coupling must agree to lowest order.

At large (and non-exceptional, see section 4.1.3) momenta where non-perturbative effects are small, dimensional analysis allows us to write,

$$\Gamma_R^{(n)}(p_i; g_R, \mu) \rightarrow \mu^{4-n} \tilde{\Gamma}_R^{(n)}\left(\frac{p_i}{\mu}; g_R\right), \quad (2.49)$$

in this situation scaling μ by $\frac{1}{s}$ is equivalent to scaling p_i by s . The Callan Symanzik equation can be solved, see [37] for example, to get,

$$\tilde{\Gamma}_R^{(n)}\left(\frac{p_i}{\mu'}; g_R(\mu')\right) = \exp\left(n \int_{g_R(\mu')}^{g_R(\mu)} \frac{\gamma(x)}{\beta(x)} dx\right) \tilde{\Gamma}_R^{(n)}\left(\frac{p_i}{\mu}; g_R(\mu)\right). \quad (2.50)$$

Aside from the canonical scaling the effect of changing μ is to change the coupling

constants and multiply by the anomalous scaling of the vertex, $e^{\int \gamma/\beta}$, similar to the effect of changing the cutoff in equation 2.7.

2.4 Composite Operator Renormalization

In many applications we do not just want to consider n-point functions but instead our physical quantities may be matrix elements of composite operators. A composite operator is a normal ordered product of fields, their derivatives and possible colour and Dirac matrices at a single point. The n-point function with the insertion of the operator $\mathcal{O}(x)$ is,

$$G_{\mathcal{O}}^{(n)}(x_1, \dots, x_n; x) = \langle \psi(x_1) \dots \psi(x_n) \bar{\psi}(0) \dots \bar{\psi}(0) \mathcal{O}(x) \rangle. \quad (2.51)$$

These Green's functions can be generated from the partition function if we modify the action by introducing a source term for the operator,

$$W[J, \Omega] = \int D\psi D\bar{\psi} \exp \left(-S_F[\psi, \bar{\psi}] - \int d^4x J(x) \bar{\psi}(x) + \bar{J}(x) \psi(x) + \Omega(x) \mathcal{O}(x) \right). \quad (2.52)$$

In perturbation theory this generates new vertices that can inject or carry off non-zero momentum from the graph. The same classification into relevant, irrelevant and marginal holds for the composite operators too. If we introduce an irrelevant operator it means the action has strayed from the renormalized trajectory and Green's functions with insertions of the operator no longer have a continuum limit. We calculate matrix elements of irrelevant operators often, especially when considering effective theories, expected to be valid only below some cutoff Λ . As long as we only want predictions at scales much less than Λ this is not a problem.

Just as corrections to couplings were generated in section 2.24, the composite operator also changes under renormalization group transformations,

$$\mathcal{O}(x) \rightarrow \mathcal{O}(x) + \delta \mathcal{O}(x). \quad (2.53)$$

It may be that the correction is proportional to the original operator, in which case it is said to be multiplicatively renormalizable, or it may be that new terms appear in the action. Unless some symmetry of the theory forbids it, a composite operator will mix with all operators of dimension lower than itself. In the case of

m operators mixing there will be m linear combinations that are multiplicatively renormalizable,

$$\mathcal{O}_i = Z_{ij} \mathcal{O}_j \rightarrow \mathcal{O}'_i = \delta_{ij} Z_{\mathcal{O}'_i} \mathcal{O}'_i, \quad (2.54)$$

see reference [50] for an example of operator mixing.

2.5 $\overline{MS}$

For perturbative calculations dimensional regularization [129] is the most convenient way of making the theory finite. All Feynman integrals are performed in $d = 4 - \epsilon$ dimensions and the limit is taken as $\epsilon \rightarrow 0$. This seems far removed from our picture above involving sharp momentum cutoffs to regulate the integrals which are now performed over all momenta in $4 - \epsilon$ dimensions but is an equally valid way to define the theory perturbatively.

We take ϕ^4 theory as an example, in $4 - \epsilon$ dimensions a scale enters in front of the coupling constants, $\lambda \rightarrow \mu^\epsilon \lambda$ and $m \rightarrow \mu^{-\epsilon/2} m$ from 2.4. It is also convenient to write the action in the form 2.42 where the parameters in the action have been split into a renormalized part and a counter-term part. We can write the bare fields, masses and coupling in terms of renormalized parameters and renormalization constants Z which contain all the scale dependence,

$$\phi_B = \phi(1 + \delta Z)^{1/2} = \phi Z_\phi^{1/2} \quad (2.55)$$

$$m_B^2 = m^2(1 + \frac{\delta m^2}{m^2}) = m^2 Z_m \quad (2.56)$$

$$\lambda_B = \lambda(1 + \delta \lambda) = \lambda Z_\lambda. \quad (2.57)$$

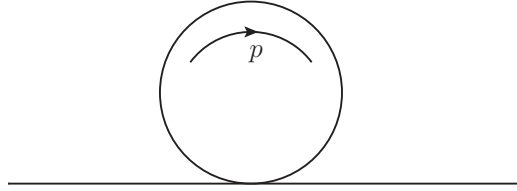


Figure 2.3: Amputated 2-point function equation 2.58.

Consider as an example of a physical quantity the amputated two point

function to leading order in λ . The $O(\lambda)$ term in the perturbative expansion is, figure 2.3 ,

$$\frac{\lambda}{2}\mu^\epsilon \int \frac{d^d p}{(2\pi)^d} \frac{1}{p^2 + m^2}, \quad (2.58)$$

the momentum integral has been kept dimensionless by including an extra factor μ^ϵ and the $\frac{1}{2}$ is a symmetry factor. Using a standard result for the volume of a sphere in d dimensions (recall we always work in Euclidean space in this thesis) this is,

$$\frac{\lambda}{2} \left(\frac{(4\pi)^{1/2}\mu}{m} \right)^\epsilon \frac{1}{16\pi^2} m^2 \Gamma\left(\frac{\epsilon}{2} - 1\right). \quad (2.59)$$

An expansion about $\epsilon = 0$ gives,

$$\frac{-\lambda m^2}{16\pi^2\epsilon} + \frac{\lambda m^2}{32\pi^2} (\gamma - 1 - \log(\frac{4\pi\mu^2}{m^2})). \quad (2.60)$$

with $\gamma = 0.57721\dots$. Thus if we choose

$$\delta m^2 = \frac{\lambda m^2}{16\pi^2\epsilon} + \frac{\lambda m^2}{16\pi^2} (-\gamma + \log(4\pi)) \quad (2.61)$$

then, to leading order in λ , the 2-point function is finite. The only divergence is in the bare mass $m^2 + \delta m^2$ but the physical mass, defined by the position of the pole in the 2-point function, is finite and given by,

$$m^2(\mu) = (m^2)^{\overline{MS}} Z_m^{\overline{MS}}(\mu) = (m^2)^{\overline{MS}} \left[1 + \frac{\lambda^{\overline{MS}}}{32\pi^2} \left(\log \frac{m^2}{\mu^2} - 1 \right) \right]. \quad (2.62)$$

All the scale dependence has been absorbed by the renormalization constant $Z_m^{\overline{MS}}(\mu)$. The choice of counter term to eliminate the poles and some finite constants that always occur in dimensional regularization defines the $\overline{MS}$ scheme.

Perturbative calculations in $\overline{MS}$ are much simpler than in other schemes but it is difficult to see how to implement $\overline{MS}$ in other regularizations, namely on the lattice. This is a problem since we often have, especially for electroweak processes, calculations done partially in perturbation theory in $\overline{MS}$ and partially via lattice QCD. Physical quantities do not depend on the scheme but the whole calculation must be done consistently in the same scheme. Using lattice perturbation theory [96] is difficult because of the many extra vertices generated and, even if the calculation can be done, the perturbative series is ill-convergent due to tadpole

graphs. In the absence of a convergent lattice perturbation theory what must be done is to calculate a conversion factor from $\overline{MS}$ to a lattice scheme. Let S_1 and S_2 denote two schemes. Then the renormalized quantities, for example the fields, will be different depending on the scheme,

$$\phi_{S_1} = Z_{S_1}^{-1/2} \phi_B \tag{2.63}$$

$$\phi_{S_2} = Z_{S_2}^{-1/2} \phi_B \tag{2.64}$$

and the quantity,

$$C^{S_1 \rightarrow S_2} = \frac{Z^{S_1}}{Z^{S_2}} \tag{2.65}$$

will relate fields defined in each scheme. This matching has been computed for many different schemes and different quantities, for example [38] and [77]. In chapter 4 we will define a scheme that we can use for lattice and perturbative QCD calculations.

Chapter 3

Lattice Methods

In this chapter we describe some useful methods and measurement techniques for lattice QCD that are relatively standard in the field and will be used in subsequent chapters.

3.1 Correlation Functions

The Euclidean correlation function of operators $\mathcal{O}_1$ and $\mathcal{O}_2^\dagger$ is

$$\langle \mathcal{O}_1 \mathcal{O}_2^\dagger \rangle = \frac{1}{Z} \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}A \exp(-S[\psi, \bar{\psi}, A]) \mathcal{O}_1[\psi, \bar{\psi}, A] \mathcal{O}_2^\dagger[\psi, \bar{\psi}, A]. \quad (3.1)$$

There is a subtlety here in that $\mathcal{O}$ on the left is meant to be a quantum operator while on the right it is a product of classical fields. We start on the left hand side and evaluate $\mathcal{O}_2$ at the origin and $\mathcal{O}_1$ at $x = (\vec{x}, t)$, $t > 0$. The time ordered vacuum expectation value on the left is then,

$$C(\vec{x}, t) \equiv \langle 0 | \mathcal{O}_1(\vec{x}, t) \mathcal{O}_2^\dagger(\vec{0}, 0) | 0 \rangle. \quad (3.2)$$

Inserting a complete set of states,

$$I = \sum_n \frac{|n\rangle \langle n|}{V}, \quad (3.3)$$

where V is the spatial volume, using the Heisenberg relation, $\mathcal{O}_1(\vec{x}, t) = e^{Ht}\mathcal{O}_1(\vec{x}, 0)e^{-Ht}$, and Fourier transforming gives,

$$C(\vec{p}, t) = \sum_{n, \vec{x}} e^{i\vec{p}\vec{x}} \frac{\langle 0 | \mathcal{O}_1(\vec{x}, 0) | n \rangle \langle n | \mathcal{O}_2^\dagger(\vec{x}, 0) | 0 \rangle}{V} e^{-E_n t}. \quad (3.4)$$

For example, if $\mathcal{O}_1 = \mathcal{O}_2 = \mathcal{O}_K$ is an operator that has the same quantum numbers as a kaon then $\langle n | \mathcal{O}_K^\dagger(\vec{0}, 0) | 0 \rangle$ vanishes for all states n that do not have these quantum numbers. The state that dominates the sum in the limit of large t is the kaon, $|n\rangle = |K\rangle$, the next is the first excited state of the kaon, $|K'\rangle$. By taking the zero momentum correlator so that $E_n = M_n$ we have,

$$C(\vec{0}, t) = C(t) = |\langle 0 | \mathcal{O}_1(\vec{0}, 0) | K \rangle|^2 e^{-M_K t} + |\langle 0 | \mathcal{O}_1(\vec{0}, 0) | K' \rangle|^2 e^{-M_{K'} t} + \dots \quad (3.5)$$

The large time behaviour of Euclidean correlation functions is an exponential decay with the exponent giving the particle masses.

The correlation function is evaluated in lattice calculations by returning to 3.1 and looking at the right hand side. After performing the fermionic path integral we get,

$$\langle \mathcal{O}_1 \mathcal{O}_2^\dagger \rangle \simeq \sum_i \langle \mathcal{O}_1 [U_i] \mathcal{O}_2^\dagger [U_i] \rangle_F. \quad (3.6)$$

The notation $\langle \dots \rangle_F$ means we have done the fermionic path integral and all the fermion fields have been contracted together according to Wick's theorem. The sum is over importance sampled gauge configurations, which approximates the $\int \mathcal{D}A$ functional integral.

Wick's theorem gives the time ordered expectation value, $\langle \rangle_F$, in terms of normal ordered products (all creation operators are moved to the left) and contractions of the fermion field. A contraction of $q(x)$ and $\bar{q}(y)$ is given by the fermion propagator $S(x, y)$. This is just the inverse of the Dirac operator,

$$D(x, z)S(z, y) = \delta(x, y). \quad (3.7)$$

For domain wall fermions in the above equation $D = \tilde{D}_{ov}$, however the equation is solved as in section 3.2.

3.1.1 Mesons

The simplest operators are the bilinears,

$$\bar{q}(x)\Gamma q'(x) \quad (3.8)$$

where Γ is some Dirac matrix and $\{q, q'\}$ are quark fields. The operators that create a kaon (K_+) or pion (π_-), for example, are

$$\mathcal{O}_{K_+}(x) = \bar{s}(x)\gamma_5 d(x) \quad (3.9)$$

$$\mathcal{O}_{\pi_-}(x) = \bar{d}(x)\gamma_5 u(x). \quad (3.10)$$

By computing the average over gauge configurations, generated with the appropriate weight, of

$$\begin{aligned} C(t) &= \sum_{\vec{x}} Tr \left[\bar{q}(\vec{x}, t) \Gamma_1 q'(\vec{x}, t) \bar{q}'(\vec{0}, 0) \Gamma_2 q(\vec{0}, 0) \right] \\ &= \sum_{\vec{x}} Tr \left[S^{(q')}(\vec{x}, t; \vec{0}, 0) \Gamma_1 S^{(q)}(\vec{0}, 0; \vec{x}, t) \Gamma_2 \right] \end{aligned} \quad (3.11)$$

and extracting the long time behaviour we get the mass of the meson corresponding to $\bar{q}(x)\Gamma q'(x)$.

The propagator $S(x, 0)$ gives the probability amplitude for creation of a quark at the source $x = (\vec{0}, 0)$ and annihilation at the sink $(\vec{x}, t)$. In fact, since the full theory is translationally invariant in space and time we can compute,

$$C(t) = \frac{1}{VT} \sum_{\vec{x}, \vec{y}, \tau} Tr \left[S^{(q')}(\vec{x}, t; \vec{y}, t + \tau) \Gamma_1 S^{(q)}(\vec{y}, t + \tau; \vec{x}, t) \Gamma_2 \right] \quad (3.12)$$

which has smaller variance and is equivalent after a gauge average, T is the time extent of the lattice.

In practice it is usually impossible to compute equation 3.12 exactly and equation 3.11 using one, or at most a few, different source points is the best we can do. In chapter 6 we will see how the volume average in equation 3.12 can be done, at least partially. For correlation functions the equation,

$$S(x, y) = \gamma_5 S^\dagger(y, x) \gamma_5 \quad (3.13)$$

is very useful. The propagation of a quark from y to x is re-interpreted as propagation of an anti-quark from x to y using the γ_5 -Hermiticity of the Dirac operator. This allows us obtain $S(0, x)$ from $S(x, 0)$ and compute the correlator $C(t)$ with only one propagator inversion.

3.1.2 Disconnected Diagrams

The operator which creates an iso-singlet is,

$$\mathcal{O}_S = \frac{1}{\sqrt{2}}(\bar{u}(x)\Gamma u(x) + \bar{d}(x)\Gamma d(x)). \quad (3.14)$$

We assume exact isospin symmetry for light quarks, as will be the case in our simulations, so

$$S = S_u = S_d. \quad (3.15)$$

Setting $\mathcal{O}_1 = \mathcal{O}_2 = \mathcal{O}_S$ and contracting we get a term,

$$C(t) = \frac{1}{VT} \sum_{\vec{x}, \vec{y}, \tau} Tr [S(x, y)\Gamma S(y, x)\Gamma] \quad (3.16)$$

as well as a term,

$$-2D(t) = -2\frac{1}{VT} \sum_{\vec{x}, \vec{y}, \tau} Tr [S(x, x)\Gamma] Tr [S(y, y)\Gamma]. \quad (3.17)$$

The two terms are shown in fig 3.1. The correlator $D(t)$ involves contractions of two fields at the same point and is very difficult to evaluate numerically. Naively it requires V point source inversions which is unfeasible on any reasonably large lattice, chapter 6 describes the approach we take in this thesis.

3.1.3 Four-Quark Operators

We briefly discuss four quark operators, specifically,

$$\mathcal{O}_{VV+AA} = (\bar{s}\gamma_\mu d)(\bar{s}\gamma_\mu d) + (\bar{s}\gamma_5\gamma_\mu d)(\bar{s}\gamma_5\gamma_\mu d). \quad (3.18)$$

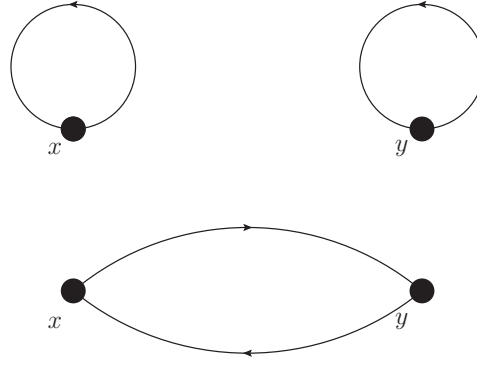


Figure 3.1: Disconnected and connected diagrams.

This is the weak effective operator that causes a transition from K^0 to $\bar{K}^0$. The parameter,

$$\epsilon_K = \frac{A(K_L \rightarrow (\pi\pi)_{I=0})}{A(K_S \rightarrow (\pi\pi)_{I=0})} \quad (3.19)$$

is experimentally known [4] and weak effective theory can relate this quantity to a perturbative function times a hadronic matrix element, for more information see [9] and references therein.

The matrix element that gives the dominant contribution is the “box diagram” describing creation of a K^0 at time t_1 , its oscillation to $\bar{K}^0$ at time t and annihilation at t_2 is

$$\langle K^0(t_1) | \mathcal{O}_{VV+AA}(t) | \bar{K}^0(t_2) \rangle. \quad (3.20)$$

Conventionally the value of,

$$B_K = \frac{\langle K^0(t_1) | \mathcal{O}_{VV+AA}(t) | \bar{K}^0(t_2) \rangle}{\frac{8}{3} f_K^2 m_K^2} \quad (3.21)$$

is reported.

We can have the s and d quarks of the K^0 at any spatial location at time t_1 and similarly for $\bar{K}^0$ at time t_2 . The vertex can be at any position on the lattice. The matrix element, where $\Gamma = \gamma_\mu(1 - \gamma_5)$ is,

$$C(t_1, t, t_2) = \sum_{\vec{y}} \langle 0 | \bar{d} \gamma_5 s(\vec{x}, t_1) \bar{s} \Gamma d(\vec{y}, t) \bar{s} \Gamma d(\vec{y}, t) \bar{d} \gamma_5 s(\vec{z}, t_2) | 0 \rangle \quad (3.22)$$

There are four possible contractions but only two inequivalent ones given by,

$$\text{Tr} [S_s(z, y) \Gamma S_d(y, z)] \text{Tr} [S_s(x, y) \Gamma S_d(y, x)] \quad (3.23)$$

$$\text{Tr} [S_s(z, y) \Gamma S_d(y, x) S_s(x, y) \Gamma S_d(y, z)] \quad (3.24)$$

these correlators are evaluated using wall sources in reference [9] and $S(y, z) = \gamma_5 S^\dagger(z, y) \gamma_5$ is used when evaluating equation 3.23.

3.2 Dirac Inversion

To calculate fermionic observables we need quark propagators. For propagators in the domain wall fermion formulation consider a point source at the origin,

$$|\eta^4(x)\rangle = \delta(x, 0) \quad (3.25)$$

here η^4 is a four dimensional vector. We write a 5d vector as a column vectors with 4d objects as entries. We construct the 5d vector

$$|\eta(x, s)\rangle = \begin{pmatrix} P_R |\eta^4(x)\rangle \\ 0 \\ \vdots \\ 0 \\ P_L |\eta^4(x)\rangle \end{pmatrix} \quad (3.26)$$

from the 4d source. Then by solving

$$D_{DWF}(x, s; y, t) |\psi(y, t)\rangle = |\eta(x, s)\rangle \quad (3.27)$$

and projecting the components of ψ ,

$$|\psi^4(x)\rangle = P_L |\psi(x, 0)\rangle + P_R |\psi(x, L_s - 1)\rangle = S(x, 0) \quad (3.28)$$

we obtain the propagator from the source at 0 to all points on the lattice.

The computational effort is in solving the matrix equation, 3.27, via the conjugate gradient (CG) algorithm [81]. This method takes a Hermitian, positive-definite matrix A and solves $Ax = b$ iteratively. Starting from a guess, x_0 and

moving along A -orthogonal search directions by an amount determined by the magnitude of the residual $r_i = b - Ax_i$ (A -orthogonality of x and y means $x^\dagger Ay = 0$) after m steps x_m is an approximate solution to the equation. For an $M \times M$ matrix, in exact arithmetic, the exact solution will be found after M steps, however, usually $m \ll M$ is sufficient for a small residual. If we have a matrix D that is not Hermitian positive definite it is simple to change the problem to one where CG can be applied (the normal equations)

$$Dx = b \rightarrow D^\dagger Dx = D^\dagger b. \quad (3.29)$$

The convergence of CG is determined by the eigenvalue spectrum of the matrix; namely the size of the smallest eigenvalue compared to the largest, called the condition number, and how dense the eigenvalues are [122].

3.2.1 Preconditioning

Preconditioning alters the equation so that the matrix we need to invert has a better condition number. There will be some overhead associated with translating the solution to the preconditioned system back into the solution of the original system, but this should be offset by the reduction in the number of iterations. Many preconditioning schemes exist, we look only at two different types of even-odd preconditioning. We assign to each site on the lattice a parity, $p = \{e, o\}$,

$$p = (x_1 + x_2 + x_3 + x_4 + x_5) \bmod(2) \quad (3.30)$$

$$p = (x_1 + x_2 + x_3 + x_4) \bmod(2) \quad (3.31)$$

for 5d or 4d preconditioning respectively. This enables a decomposition of D_{DWF} into different parity components,

$$\begin{aligned} D_{DWF} &= \begin{pmatrix} M_{ee} & M_{eo} \\ M_{oe} & M_{oo} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ M_{oe}M_{ee}^{-1} & 1 \end{pmatrix} \begin{pmatrix} M_{ee} & 0 \\ 0 & D_{oo} \end{pmatrix} \begin{pmatrix} 1 & M_{ee}^{-1}M_{eo} \\ 0 & 1 \end{pmatrix} \\ &= LDU \end{aligned} \quad (3.32)$$

where we have Schur decomposed the matrix into its LDU factorization and

$$D_{oo} = M_{oo} - M_{oe}M_{ee}^{-1}M_{eo}. \quad (3.33)$$

Now equation 3.27 becomes,

$$\begin{aligned} D_{DWF}\psi &= \eta \\ LDU\psi &= \eta \\ DU\psi &= L^{-1}\eta \\ \begin{pmatrix} M_{ee} & M_{eo} \\ 0 & D_{oo} \end{pmatrix} \psi &= \begin{pmatrix} 1 & 0 \\ -M_{oe}M_{ee}^{-1} & 1 \end{pmatrix} \eta \end{aligned} \quad (3.34)$$

We decompose the vectors into even and odd subspaces,

$$\begin{aligned} \begin{pmatrix} M_{ee} & M_{eo} \\ 0 & D_{oo} \end{pmatrix} \begin{pmatrix} \psi_e \\ \psi_o \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ -M_{oe}M_{ee}^{-1} & 1 \end{pmatrix} \begin{pmatrix} \eta_e \\ \eta_o \end{pmatrix} \\ \begin{pmatrix} M_{ee}\psi_e + M_{eo}\psi_o \\ D_{oo}\psi_o \end{pmatrix} &= \begin{pmatrix} \eta_e \\ \eta_o - M_{oe}M_{ee}^{-1}\eta_e \end{pmatrix} = \begin{pmatrix} \eta_e \\ \eta'_o \end{pmatrix}. \end{aligned} \quad (3.35)$$

From this we find that, by solving the equation,

$$D_{oo}^\dagger D_{oo} \psi_o = D_{oo}^\dagger \eta'_o \quad (3.36)$$

with

$$\eta'_o = \eta_o - M_{oe}M_{ee}^{-1}\eta_e \quad (3.37)$$

we can construct

$$\psi_e = M_{ee}^{-1}(\eta_e - M_{eo}\psi_o) \quad (3.38)$$

to obtain the full solution. The matrix $D_{oo}^\dagger D_{oo}$ has a smaller condition number than $D_{DWF}^\dagger D_{DWF}$ and this trick reduces the cost of CG by a factor of around 2 or 3.

We will mainly be working with $4d$ even-odd preconditioning. In this case only the Wilson term couples sites of opposite parity so $M_{oe} = (D_W)_{oe}$ and similarly

$M_{eo} = (D_W)_{eo}$. The remaining diagonal terms give M_{ee} and M_{oo} , both equal to,

$$M_{ee} = \begin{pmatrix} K & -P_L & 0 & \dots & 0 & mP_R \\ -P_R & K & -P_L & \dots & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & \ddots & -P_L \\ mP_L & 0 & \dots & 0 & -P_R & K \end{pmatrix}. \quad (3.39)$$

with $K = 5 - M_5$. We also need the inverse of this which is,

$$M_{ee}^{-1} = d(AP_R + A^T P_L) \quad (3.40)$$

$$A = \begin{pmatrix} K^{L_s-1} & -m & -mK & \dots & -mK^{L_s-2} \\ K^{L_s-2} & K^{L_s-1} & -m & \dots & -mK^{L_s-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ K^0 & K^1 & K^2 & \dots & K^{L_s-1} \end{pmatrix}. \quad (3.41)$$

and

$$d = \frac{1}{m + K^{L_s}} \quad (3.42)$$

3.3 Gauge Fixing

The vertex functions that are required for Rome Southampton renormalization are not gauge invariant. Elitzur's theorem [59] implies that gauge non-invariant observables vanish under a gauge average. However, it is often useful to evaluate observables that are not gauge invariant. Such quantities must be computed in a fixed gauge. This is often the case if we want to compare lattice calculations with continuum (perturbative) calculations done in a particular gauge. On the lattice the gauge of choice is usually the Landau gauge, which is fixed by the condition that,

$$\partial_\mu A_\mu = 0, \quad (3.43)$$

which is the condition for the extrimization of some functional of A . This functional is [132],

$$W[A] = \sum_{\mu} \int d^4x \text{tr}[A_{\mu}(x)^2]. \quad (3.44)$$

In terms of the link matrices, minimizing this functional is equivalent to minimizing the functional,

$$W[U] = - \sum_x \sum_{\mu} \text{tr}[U_{\mu}(x) + U_{\mu}(x)^{\dagger}], \quad (3.45)$$

which can be seen by expanding $U_{\mu}(x) = \exp(iaA_{\mu}(x))$.

We seek a gauge transformation $\Omega(x)$ such that

$$U_{\mu}(x) \rightarrow U_{\mu}^{\Omega}(x) = \Omega(x)U_{\mu}(x)\Omega(x + \mu)^{\dagger} \quad (3.46)$$

and $W[U_{\mu}^{\Omega}]$ is minimized. This minimization is done by choosing a gauge transformation matrix $g(x)$ at each x to minimize W and then repeating this at each lattice site for some number of iterations through the whole lattice until the value of the functional W stops changing. Such a method can only find local minima and there remains the so-called Gribov ambiguity that stems from our inability to find the global minimum. In this thesis where gauge fixing is used we will always fix the gauge to sufficient accuracy that any uncertainty coming from the Gribov noise is well below the level of our statistical error.

3.4 Fitting

Suppose we want to extract the mass of a particle from equation 3.5. For concreteness consider the function,

$$Y(t; A, M) = A \exp(-Mt) \quad (3.47)$$

and the data that is supposed to be described by this,

$$\{y_i(1), y_i(2), \dots, y_i(T)\} \quad (3.48)$$

where $i = \{1, \dots, N\}$ labels the gauge configuration. We want to estimate the mean and variance of y at each time step, fit this data to $Y(t; A, M)$ and extract the mass. Simply adding up the data on different configurations will introduce bias. This is because new gauge configurations are generated by updating old ones so configurations nearby in molecular dynamics “time” are correlated. The similarity between nearby configurations can lead to underestimates of the true variance. Typically measurements are only performed on a subset of gauge configurations, for example every tenth one, but this alone may not eliminate correlations. Measurements nearby in molecular dynamics time can be binned (averaged) together, giving a smaller set of measurements. The analysis is performed on the smaller set of binned data. By progressively increasing the number of samples in each bin the correlations should decrease until a change in bin width no longer affects the error.

The propagation of statistical error through the calculation is done using re-sampling techniques, we use either jackknife or bootstrap. Bootstrap resampling takes the original data $\{y_i(t)\}$ and draws N samples, with replacement, at random to make an estimate of the average $\bar{y}(t)$ from these samples. This is done N_{boot} times where $N_{boot} \sim O(300)$ resulting in a distribution of averages, $\{\bar{y}_j(t)\}$, with $j = \{1, \dots, N_{boot}\}$. The variance is estimated by sorting $\{\bar{y}_j(t)\}$ and taking half the difference between the upper and lower bounds of a confidence interval around the centre of the distribution. Assuming Gaussian statistics a 68% confidence interval gives a one- σ error.

Jackknife resampling discards one measurement and calculates the average from those remaining, doing this for all N samples gives the data $\bar{y}_j(t)$ $j = \{1, \dots, N\}$ and the error on the mean is estimated as [52],

$$\sigma^2(t) = \frac{N-1}{N} \sum_{j=1}^N (\bar{y}_j(t) - \langle \bar{y}(t) \rangle)^2 \quad (3.49)$$

where $\langle \bar{y}(t) \rangle$ is the average of $\{\bar{y}_j(t)\}$.

The fit is performed by minimizing the function,

$$\chi_j^2 = \sum_{t_1, t_2} (\bar{y}_j(t_1) - Y(t_1; A, M)) C^{-1}(t_1, t_2) (\bar{y}_j(t_2) - Y(t_2; A, M)) \quad (3.50)$$

with respect to A and M . The function is minimized for every resampled data

point giving us a distribution of A and M from which we can estimate the average and standard error. The matrix C is the covariance matrix which is,

$$C(t_1, t_2) = \frac{1}{N} \sum_i^N \sum_j^N (\bar{y}_j(t_1) - \langle \bar{y}(t_1) \rangle) (\bar{y}_j(t_2) - \langle \bar{y}(t_2) \rangle) \quad (3.51)$$

and is calculated from the unresampled data. It is the same for each fit over the resampled distribution which is known as using a frozen covariance matrix. Setting the covariance matrix to the unit matrix will perform an uncorrelated fit. Although it is important to control correlations and we make every effort to do so, for very noisy data the error induced by correlations between the gauge fields is swamped by the noise in the signal. In such cases, including more gauge configurations closer together in molecular dynamics time, and therefore highly correlated with each other, will be necessary to reduce the error which is large enough to overwhelm the effect of correlations between gauge configurations.

3.4.1 Black Box Fits

From equation 3.5 we can see that the leading correction to the simple exponential decay of the correlation function is another exponential with larger mass. At large times this will be small, but, on a finite lattice, there may not be enough time for the excited state signal to decay completely. It is possible to fit the data to multiple exponentials to account for this. A more sophisticated approach is the “black box” method [61]. Write the multi-exponential decay

$$Y(t) = \sum_{k=1}^K A_k \exp(-M_k)t = \sum_{k=1}^K A_k \alpha_k^t \quad (3.52)$$

where $\alpha_k = \exp(-M_k)$. If there is only one exponential then define the effective mass

$$m_{eff}(t) = \log \left(\frac{Y(t)}{Y(t+1)} \right) = -\log(\alpha_1) = M_1. \quad (3.53)$$

By plotting this and fitting a constant where the data shows a plateau we can determine the mass. The case of two exponentials can be handled similarly by taking four datapoints and generally for K exponentials we need $N = 2K$ datapoints. For overdetermined systems $N > 2K$ the following method, called

linear prediction, can be used. Define the polynomials,

$$p(\alpha) = \prod_{i=1}^K (\alpha - \alpha_i) = \sum_{i=0}^K p_i \alpha^{K-i}, \quad p_0 = 1 \quad (3.54)$$

since $p(\alpha_k) = 0$ then

$$\alpha_k^m = - \sum_{i=1}^K p_i \alpha_k^{m-i}, \quad m \geq K \quad (3.55)$$

which on substitution into 3.52 gives, for $t \geq K$,

$$Y(t) = - \sum_{k=1}^K A_k \sum_{i=1}^K p_i \alpha_k^{t-i} = - \sum_{k=1}^K p_k Y(t-k). \quad (3.56)$$

Solving for the coefficients p_k and subsequently obtaining the roots of the polynomial $p(\alpha)$ gives α_k and hence the masses. In this way earlier time data, which is usually more accurate, may be used to extract effective masses.

3.5 Sommer Scale

A Wilson loop is given by the trace of the product of the links around a closed loop L ,

$$W(\vec{r}, t) = Tr \left[\prod_{x, \mu \in L} U_\mu(x) \right]. \quad (3.57)$$

Wilson loops are gauge invariant observables, physically they give the potential between an infinitely massive quark anti-quark pair separated by $\vec{r}$,

$$\langle W(\vec{r}, t) \rangle \propto e^{-V(\vec{r})t}. \quad (3.58)$$

Non-planar Wilson loops can be measured to estimate of the potential at separations other than integer multiples of the lattice spacing, however, on the lattice we do not have rotational invariance and different directions not related by ninety degree rotation or reflection are not equivalent. Thus non-planar Wilson loops will have different lattice artefacts than planar ones. This will only be an issue if the statistical precision is enough to resolve the difference.

By measuring expectation values of Wilson loops we can extract the potential which is expected to be of the form,

$$V(r) = A + \frac{B}{r} + \sigma r, \quad (3.59)$$

$r = |\vec{r}|$. There is a linearly rising part due to the confining nature of the force and a short distance Coulomb term. With dynamical fermions the form of the potential given above is valid only at low energies. When the energy between the quarks is great enough to create a quark anti-quark pair out of the vacuum the flux tube joining the two quarks will break, creating two mesons. In our simulations we do not expect to reach high enough energies to observe string breaking effects. The string tension, σ , is a physical scale that can be measured in lattice simulations and from the experimental effective inter-quark potential. However extracting this constant relies on long distance behaviour of the force which is hard to calculate accurately on the lattice.

Instead one defines a scale, $r(c)$, called the Sommer scale [125] by the equation,

$$r^2(c)F(r(c)) = c \quad (3.60)$$

where F is the force derived from the potential and c is a constant chosen such that the distance $r(c)$ is small enough to measure accurately on the lattice. Usually $c = 1.65$ and this is “the” Sommer scale r_0 , although other values eg. $c = 1$, $c = 0.65$ are often used. For $c = 1.65$, by comparing with experimental effective quark potentials, $r(c) \simeq 0.5 fm$. This approximate relation means that setting the lattice scale using r_0 is not very accurate, however, since the lattice determination of r_0 is possible to do accurately it is useful for relative scale setting, measuring the ratio of two different lattice spacings.

The force can be naively defined from the potential by a finite difference,

$$F(r - \frac{a}{2}) = \frac{V(r) - V(r - a)}{a}, \quad (3.61)$$

however, it is found to greatly reduce discretization effects [106], if we define a tree level improved force,

$$F(r_I) = \frac{V(r) - V(r - a)}{a}. \quad (3.62)$$

The distance r_I is calculated for r on the right hand side by solving

$$\frac{1}{4\pi r_I^2} = \frac{G(r-a, 0, 0) - G(r, 0, 0)}{a} \quad (3.63)$$

$$G(\vec{r}) = \frac{1}{a} \int_{-\pi}^{\pi} \frac{d^3 k}{(2\pi)^3} \frac{\prod_{j=1}^3 \cos(r_j k_j / a)}{4 \left(\sum_{j=1}^3 \sin^2(k_j/2) - 4c_1 \sum_{j=1}^3 \sin^4(k_j/2) \right)}. \quad (3.64)$$

We then have the force at r_I calculated from the potential at r and $r-a$. With this definition the force at r_I is $\frac{4}{3} \frac{g_0^2}{4\pi r_I^2}$, removing the $O((a/r)^2)$ lattice artefacts from the naive choice. G is the scalar lattice propagator and depends on the action. The co-efficient c_1 is an improvement co-efficient for the gauge action, see equation 1.37, and we evaluate G for the Iwasaki action used in this thesis later, in section 4.6.2.

Chapter 4

Step Scaling with Rome-Southampton Renormalization

The non-perturbative renormalization (NPR) of QCD matrix elements is an important topic for combining lattice QCD calculations and phenomenology. In this chapter we will describe the regularization invariant schemes which are used to intermediate between lattice data and perturbative calculations. We will implement the different schemes to renormalize several phenomenologically important quantities, the quark field, quark mass, tensor current and the kaon bag parameter, B_K .

Common to all our schemes is the window condition on the simulated momenta

$$\Lambda_{QCD}^2 \ll \mu^2 \ll \left(\frac{\pi}{a}\right)^2. \quad (4.1)$$

The degree to which we satisfy these inequalities is a source of uncertainty in many calculations and leads to an entanglement of perturbative truncation errors and lattice artefacts. In this chapter we describe a method to open the window by continuum extrapolating to deal with lattice artefacts and then raising the scale, μ , non-perturbatively to bring us into a regime where perturbation theory can be safely applied. In this way the perturbative errors are cleanly separated from the lattice artefacts and we can, for the first time, rigorously account for both.

4.1 Rome-Southampton Renormalization

4.1.1 RI-MOM

We renormalize on the lattice using the Rome-Southampton method [103] and its variants [127]. This means using schemes called RI-MOM schemes (regularization invariant momentum schemes) that can be implemented easily on the lattice or in dimensional regularization. It might seem more natural to compute hadron masses and fix their values to what is experimentally measured in order to define the renormalization scheme. This can be done, but in such schemes it is hard to match perturbative calculations with lattice results. RI-MOM schemes facilitate this matching.

Denote the renormalized propagator by S_R , this is the propagator constructed with the renormalized field q_R and renormalized mass m_R ,

$$q_R(x) = q_B(x) Z_q^{1/2} \quad (4.2)$$

$$m_R = m_B Z_m = (m_f + m_{res}) Z_m \quad (4.3)$$

$$S_R(p, m_R) = Z_q S_B(p, m_B)|_{m_B=m_R/Z_m}. \quad (4.4)$$

The renormalization constants in this scheme are fixed by the requirement, see [24], that

$$\lim_{m_R \rightarrow 0} \frac{1}{12m_R} \text{Tr} [S_R^{-1}(p)]_{p^2=\mu^2} = 1 \quad (4.5)$$

$$\lim_{m_R \rightarrow 0} \frac{1}{48} \text{Tr} \left[\gamma_\mu \frac{\partial S_R^{-1}(p)}{\partial p_\mu} \right]_{p^2=\mu^2} = 1 \quad (4.6)$$

this defines the RI-MOM scheme. Changing equation 4.6 to,

$$\lim_{m_R \rightarrow 0} \frac{1}{12p^2} \text{Tr} [S_R^{-1}(p) \not{p}]_{p^2=\mu^2} = 1 \quad (4.7)$$

defines the closely related RI'-MOM scheme, see [38], [77]. The tree level propagator $\not{p} + m$ satisfies these equations, so the renormalization conditions are equivalent to imposing that the full renormalized quark propagator is equal to its tree level value.

The renormalization of the composite operator $\mathcal{O}$ is performed as in section 2.4 by inserting the operator into the partition function with its own source,

generating a new interaction vertex. The bilinear operators,

$$S(x) = \bar{q}(x)q'(x) \quad (4.8)$$

$$P(x) = \bar{q}(x)\gamma_5 q'(x) \quad (4.9)$$

$$V(x) = \bar{q}(x)\gamma_\mu q'(x) \quad (4.10)$$

$$A(x) = \bar{q}(x)\gamma_\mu\gamma_5 q'(x) \quad (4.11)$$

$$T(x) = \bar{q}(x)\sigma_{\mu\nu} q'(x) \quad (4.12)$$

renormalize multiplicatively because there are no other operators with lower dimension and the correct symmetry to mix with them. We calculate their expectation values on the lattice, $G_{\mathcal{O}}(p)$, and then use the inverse propagator $S^{-1}(p)$ to amputate the external legs giving

$$\Pi_{\mathcal{O}}(p) = S^{-1}(p)G_{\mathcal{O}}(p)S^{-1}(p). \quad (4.13)$$

These functions are not gauge invariant so we first fix the gauge to the Landau gauge, see 3.3. All Rome Southampton vertex functions will be assumed fixed to Landau gauge without further comment. The RI-MOM and RI'-MOM schemes choose the momenta of the external propagators equal so that the momentum injection at the vertex is zero, figure 4.1.

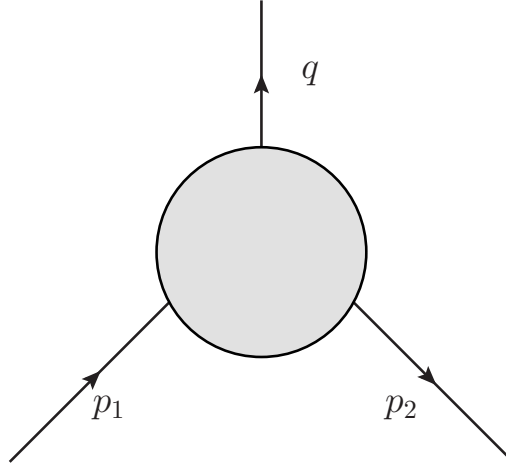


Figure 4.1: Operator vertex. For the RI and RI'-MOM schemes $p_1 = p_2 = p$, $p^2 = \mu^2$ and $q = (p_1 - p_2) = 0$.

It is useful to consider the normalized trace of the amputated vertex function with a projector $\mathbf{P}$,

$$\Lambda_{\mathcal{O}} = \frac{1}{12} \text{Tr} [\mathbf{P}_{\mathcal{O}} \Pi_{\mathcal{O}}]. \quad (4.14)$$

the projector is chosen so that, at tree level, this is equal to one. For the bilinears,

$$\mathbf{P}_S = 1 \quad (4.15)$$

$$\mathbf{P}_P = \gamma_5 \quad (4.16)$$

$$\mathbf{P}_V = \frac{1}{4} \gamma_\mu \quad (4.17)$$

$$\mathbf{P}_A = \frac{1}{4} \gamma_\mu \gamma_5 \quad (4.18)$$

$$\mathbf{P}_T = \frac{1}{12} \sigma_{\mu\nu} \quad (4.19)$$

The renormalized amputated vertex functions of bilinears are related to the bare ones by

$$\Lambda_{\mathcal{O},R} = \frac{Z_{\mathcal{O}}}{Z_q} \Lambda_{\mathcal{O},B}, \quad (4.20)$$

the Z_q factors come from the inverse propagators required for the external leg amputation. The renormalization condition on the projected vertex functions is,

$$\Lambda_{\mathcal{O}R} = 1 \quad (4.21)$$

at $m_R = 0$.

The Ward-Takahashi identities express the symmetries of QCD as relations between vertex functions and propagators. Thus we must show that the conditions on the propagators, equations 4.5 and 4.6, are consistent with the conditions on the vertex functions. For the conserved vector and conserved axial currents the Ward-Takahashi identities due to chiral symmetry are

$$q_\mu \Pi_{V,B}^\mu(p_2, p_1) = S_B^{-1}(p_2) - S_B^{-1}(p_1) \quad (4.22)$$

$$-q_\mu \Pi_{A,B}^\mu(p_2, p_1) = 2m_B \Pi_{P,B}(p_1, p_2) - \gamma_5 S_B^{-1}(p_1) - S_B^{-1}(p_2) \gamma_5 \quad (4.23)$$

(compare to section 1.4.1). First we check that the renormalization conditions of equation 4.21 applied to the local vector current are consistent with the RI-MOM conditions. In the chiral limit differentiate both sides of 4.22 with respect to q at

$p_1 = p - \frac{q}{2}$, $p_2 = p + \frac{q}{2}$ and take $q = 0$ to obtain,

$$Z_V \left(\Pi_{V,B}^\mu(p) + q_\mu \frac{\partial}{\partial q_\nu} \Pi_{V,B}^\mu(p + \frac{q}{2}, p - \frac{q}{2}) \Big|_{q=0} \right) = \frac{\partial S_B^{-1}(p)}{\partial p_\mu} \quad (4.24)$$

the second term on the left vanishes at $q = 0$. After we multiply by $\frac{1}{12} \mathbf{P}_V$ and trace over colour and Dirac indices the left hand side is

$$\frac{1}{12} Z_V \text{Tr} [\mathbf{P}_V \Pi_{V,B}^\mu(p)] = Z_V \Lambda_{V,B} \quad (4.25)$$

and the right,

$$\frac{1}{48} \text{Tr} \left[\gamma_\mu \frac{\partial S_B^{-1}(p)}{\partial p_\mu} \right] = \frac{1}{48} \text{Tr} \left[\gamma_\mu \frac{\partial Z_q S_R^{-1}(p)}{\partial p_\mu} \right] = Z_q. \quad (4.26)$$

using equation 4.6 in the last step. We then have,

$$\frac{Z_V}{Z_q} \Lambda_{V,B} = \Lambda_{V,R} = 1 \quad (4.27)$$

consistent with equation 4.21 for the vector vertex. The axial vertex is similar but the analogue of equation 4.24 only works at high μ , in this case the $q_\mu \frac{\partial}{\partial q_\nu} \Pi_{A,B}^\mu$ factor doesn't vanish at $q = 0$ for small μ and the renormalization condition becomes consistent only at high energy.

Taking the axial Ward identity, equation 4.23, at $p_1 = p - \frac{q}{2}$, $p_2 = p + \frac{q}{2}$ and $q = 0$ gives

$$\begin{aligned} 2m_B \Pi_{P,B}(p_1, p_2) &= i\gamma_5 S_B^{-1}(p_1) + iS_B^{-1}(p_2) \gamma_5 \\ 2 \frac{m_R}{Z_m} \frac{Z_q}{Z_P} \Pi_{P,R}(p_1, p_2) &= Z_q (i\gamma_5 S_B^{-1}(p_1) + iS_B^{-1}(p_2) \gamma_5). \end{aligned} \quad (4.28)$$

Multiply by $\mathbf{P}_P$ and tracing,

$$2 \frac{1}{Z_P} \Lambda_{P,R} = 2Z_m \frac{1}{12m_R} \text{Tr} [S_R^{-1}(p)]. \quad (4.29)$$

then taking the chiral limit gives,

$$Z_m = \frac{1}{Z_P}. \quad (4.30)$$

Chiral symmetry implies that $Z_P = Z_S$ (to $O(m_{res}a)$) so $Z_m = \frac{1}{Z_S}$ also. An explicit analysis [24] shows that Λ_S suffers more from spontaneous chiral symmetry breaking and gets large contributions from zero modes that are only suppressed at very high p^2 .

The conserved vector and axial currents are not renormalized (for exact chiral symmetry),

$$Z_V = Z_A = 1. \quad (4.31)$$

The local currents are more convenient to work with and satisfy,

$$Z_V = Z_A \quad (4.32)$$

where the actual value is calculated as in section 1.4.1, see equation 1.70. We can use the conserved currents to give two different definitions of Z_q ,

$$Z_q^{(V,A)} = \Lambda_{(V,A),B} Z_{(V,A)} \quad (4.33)$$

these will differ at low energy due to spontaneous chiral symmetry breaking. Alternatively Z_q could be obtained directly from the propagator and condition 4.6. In practice it is better to exploit the fact that Z_A can be computed very accurately by the methods of section 1.4.1 and use Λ_A to divide out Z_q ,

$$Z_O = Z_A \frac{\Lambda_{A,B}}{\Lambda_{O,B}}. \quad (4.34)$$

then eliminate Z_A with the previously computed value $\frac{Z_A}{Z_A}$ where $Z_A = 1$ is assumed to hold. This procedure has been adopted in previous work [8]. However following the discussion of reference [10], appendix B, $Z_A = 1$ is violated at order $m_{res}a$. The same work shows that, to this order, $Z_V = 1$. Therefore, if we calculate $\frac{Z_V}{Z_V}$ and use this to remove the factors of Z_V and/or Z_A we will not have to account for the $\sim 1\%$ error that comes from assuming $Z_A = 1$. On our smaller 16^3 lattices $\frac{Z_V}{Z_V}$ has not been calculated and we use $\frac{Z_A}{Z_A}$, this introduces some error but the renormalization constants on these small volumes are for illustration only. Our intention for small volume lattices is to use them to compute step scaling functions, section 4.5. This involves a ratio that cancels Z_A so these considerations are irrelevant for step scaling. On the larger volumes (24^3 and 32^3) we use $\frac{Z_V}{Z_V}$.

4.1.2 RI-SMOM

The RI-MOM and RI'-MOM conditions have been used to renormalize operators in lattice QCD for many years. However it has been noted [8] that the choice of momentum routing in the RI-MOM vertices is not optimal for minimizing non-perturbative effects. These arise when we have an internal line with vanishing momentum. This means the vertices can couple to the low energy regime of QCD, getting contributions from the quark condensate, $\langle \bar{q}q \rangle$, for example. The RI-MOM condition is defined at the exceptional point in momentum space,

$$p_1^2 = p_2^2 = \mu^2 \quad (p_1 - p_2)^2 = 0 \quad (4.35)$$

shown in figure 4.2(a). With an exceptional momentum configuration it is possible for a gluon to transfer all the momentum p through the vertex and leave only very low energy modes in the upper part of the graph. Because we need at least one gluon propagator this effect falls with momentum like $\frac{1}{p^2}$. On the lattice we are restricted to using momenta less than $\approx \frac{\pi}{a}$ so this rate of fall-off may not be fast enough.

Choosing a different kinematic configuration, where the leg momenta form an equilateral triangle in momentum space,

$$p_1^2 = p_2^2 = (p_1 - p_2)^2 = \mu^2 \quad (4.36)$$

gives a non-exceptional momentum point (by non-exceptional we will always mean this, symmetric configuration see figure 4.2(b)). With these kinematics a single high energy gluon is forbidden by momentum conservation, in order to couple to soft sub-graphs at least three gluon propagators are needed. This leads to a suppression of non-perturbative effects by a factor of $\frac{1}{p^6}$ [8].

Analogously to the RI'-MOM schemes the RI-SMOM scheme is defined by, [127], [78],

$$\lim_{m_R \rightarrow 0} \frac{1}{12m_R} \left\{ \text{Tr} [S_R^{-1}(p)]_{p^2=-\mu^2} - \frac{1}{2} \text{Tr} [q_\mu \Pi_{AR}^\mu(p_1, p_2) \gamma_5]_{sym} \right\} = 1 \quad (4.37)$$

$$\lim_{m_R \rightarrow 0} \frac{1}{12p^2} \text{Tr} [S_R^{-1}(p) \not{p}]_{sym} = 1. \quad (4.38)$$

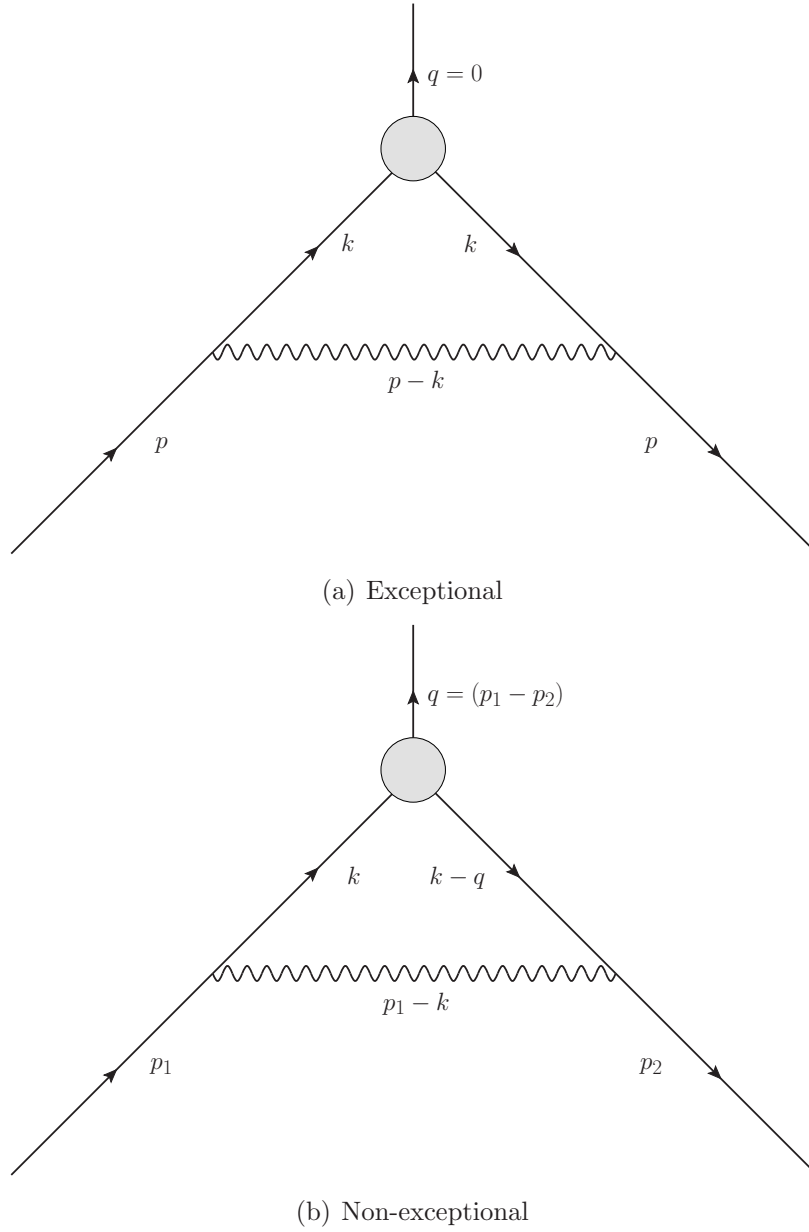


Figure 4.2: Exceptional and non-Exceptional vertex functions with one gluon exchange. In (b) it is kinematically forbidden for k to vanish for non-zero p_1 , p_2 , q satisfying the non-exceptional constraint.

The conditions on the local vector and axial vertex functions also change,

$$\lim_{m_R \rightarrow 0} \frac{1}{12q^2} \text{Tr} [q_\mu \Pi_{V,R}^\mu(p_1, p_2) \not{q}]_{sym} = 1 \quad (4.39)$$

$$\lim_{m_R \rightarrow 0} \frac{1}{12q^2} \text{Tr} [q_\mu \Pi_{A,R}^\mu(p_1, p_2) \gamma_5 \not{q}]_{sym} = 1 \quad (4.40)$$

so,

$$\mathbf{P}_V = \frac{1}{q^2} \not{q} q_\mu \quad (4.41)$$

$$\mathbf{P}_A = \frac{1}{q^2} \gamma_5 \not{q} q_\mu. \quad (4.42)$$

the pseudoscalar, scalar and tensor projectors are the same and the momenta are always in the symmetric configuration 4.36.

In this scheme, in the isospin limit, the vector Ward identity 4.22 implies,

$$\frac{1}{12q^2} \text{Tr} [q_\mu \Pi_{V,B}^\mu(p_1, p_2) \not{q}]_{sym} = \frac{1}{12q^2} \text{Tr} [S_B^{-1}(p_2) \not{q} - S_B^{-1}(p_1) \not{q}]_{sym} = -\frac{1}{12q^2} \text{Tr} [S_B^{-1}(q) \not{q}]_{sym}. \quad (4.43)$$

where the last equation follows from expanding the inverse propagator,

$$S_B^{-1}(p) \not{p} = \Sigma(p^2) p^2 + \tilde{S}(p^2) \not{p} \quad (4.44)$$

and the symmetric condition on the momenta 4.36. Then all the terms with a $\not{p}$ are zero under the trace. The axial Ward identity 4.23 implies,

$$\lim_{m_R \rightarrow 0} \frac{1}{12q^2} \text{Tr} [q_\mu \Pi_{A,B}^\mu(p_1, p_2) \gamma_5 \not{q}]_{sym} = \lim_{m_R \rightarrow 0} -\frac{1}{12q^2} \text{Tr} [S_B^{-1}(q) \not{q}]_{sym}. \quad (4.45)$$

and hence,

$$\begin{aligned} & \lim_{m_R \rightarrow 0} \frac{Z_q}{Z_V} \frac{1}{12q^2} \text{Tr} [q_\mu \Pi_{V,B}^\mu(p_1, p_2) \not{q}]_{sym} \\ &= \lim_{m_R \rightarrow 0} \frac{Z_q}{Z_A} \frac{1}{12q^2} \text{Tr} [q_\mu \Pi_{A,B}^\mu(p_1, p_2) \gamma_5 \not{q}]_{sym} \end{aligned} \quad (4.46)$$

where we have used the renormalization conditions. Thus $Z_A = Z_V$ irrespective of momentum scale μ in a chirally symmetric theory. The axial and vector vertices will still be different due to non-perturbative effects but these are suppressed by

the choice of kinematics. We also have $Z_m = \frac{1}{Z_P}$ as before and $Z_S = Z_P$. The scalar and pseudoscalar vertices are different due to spontaneous chiral symmetry breaking effects but these are suppressed compared to the RI-MOM scheme.

There is a related scheme, again using non-exceptional kinematics, called RI-SMOM $_{\gamma_\mu}$ [127], where the projectors for the vector and axial vertices are the same as in the RI-MOM scheme. This means the conditions on S_R have to be modified and the $\gamma_\mu \frac{\partial}{\partial p^\mu}$ term comes back. The conditions on the propagator are,

$$\lim_{m_R \rightarrow 0} \frac{1}{12m_R} \left\{ \text{Tr} [S_R^{-1}(p)]_{p^2=-\mu^2} - \frac{1}{2} \text{Tr} [q_\mu \Pi_{A,R}^\mu(p_1, p_2) \gamma_5]_{sym} \right\} = 1 \quad (4.47)$$

$$\lim_{m_R \rightarrow 0} \frac{1}{48} \left\{ \text{Tr} \left[\gamma^\mu \frac{\partial S_R^{-1}(p)}{\partial p^\mu} \right]_{p^2=\mu^2} + \text{Tr} \left[q_\mu \gamma^\nu \frac{\partial}{\partial q^\nu} \Pi_{V,R}^\mu \gamma_5 \right]_{sym} \right\} = -1 \quad (4.48)$$

and the projectors for the vector and axial vertices change,

$$\mathbf{P}_V = \frac{1}{4} \gamma_\mu \quad (4.49)$$

$$\mathbf{P}_A = \frac{1}{4} \gamma_\mu \gamma_5, \quad (4.50)$$

they are as in the RI-MOM scheme. There seems no a priori reason to prefer either scheme so we will do calculations in both and check after the fact which is better with respect to lattice artefacts and perturbative convergence.

4.1.3 Matching

The usefulness of the RI-MOM and related schemes is due to their generality. They are not tied to the details of the regulator like $\overline{MS}$ or lattice schemes based on hadron masses. This means the same renormalization constants can be calculated in a lattice regularization, which correctly includes non perturbative effects at low energy and in dimensional regularization where the momenta can be taken arbitrarily high. The perturbative calculation using dimensional regularization can be used to match the RI scheme with $\overline{MS}$ by calculating,

$$C_O^{RI \rightarrow \overline{MS}}(\mu) = \frac{Z_O^{\overline{MS}}(\mu)}{Z_O^{RI}(\mu)}. \quad (4.51)$$

With both renormalization constants in the same scheme we can consider the ratios,

$$\sigma_{\mathcal{O}}(\mu, \mu') = \frac{Z_{\mathcal{O}}(\mu')}{Z_{\mathcal{O}}(\mu)} \quad (4.52)$$

which relate renormalization constants at different scales. Equation 2.50 shows how these are related to the perturbatively calculated gamma and beta functions in the high energy limit.

Take the quark mass for example. The renormalization group equation is,

$$\mu^2 \frac{dm}{d\mu^2} = -m\gamma(\alpha(\mu)) = -m \left(\gamma_m^{(0)} \frac{\alpha}{\pi} + \gamma_m^{(1)} \left(\frac{\alpha}{\pi} \right)^2 + O(\alpha^3) \right) \quad (4.53)$$

where we have written the strong coupling constant $\alpha = \frac{g^2}{4\pi}$. The anomalous dimension γ_m has been perturbatively calculated to four loop order in the $\overline{MS}$ scheme [38]. Then since

$$m^{\overline{MS}} = C_m m^{RI} \quad (4.54)$$

$$\gamma^{RI} = \gamma^{\overline{MS}} - \beta^{\overline{MS}} \frac{\partial \ln C_m}{\partial \alpha}. \quad (4.55)$$

The anomalous dimension tells us how the mass evolves as we change the momentum scale μ by solving equation 4.53,

$$\frac{m(\mu')}{m(\mu)} = \frac{Z_m(\mu')}{Z_m(\mu)} = \exp \left(\int_{\alpha(\mu)}^{\alpha(\mu')} \frac{\gamma(x)}{\beta(x)} dx \right). \quad (4.56)$$

This means we can calculate a quantity in lattice QCD at low energy in an RI scheme, evolve it to high energy scales and convert it to $\overline{MS}$. However this is only true if the conversion factors, $C(\alpha(\mu))$, are known accurately which requires high order perturbation theory. The higher the energy the smaller α and the more accurate the perturbative expansion is. The requirement to have accurate perturbative results is usually stated as $\Lambda_{QCD}^2 \ll \mu^2$ where Λ_{QCD} is the strong coupling scale beyond which the perturbative expansion breaks down. On a lattice we have a natural momentum cutoff, $\frac{\pi}{a}$, and if we do not take a continuum limit we have to have $\mu^2 \ll \left(\frac{\pi}{a} \right)^2$ to be close to the renormalized trajectory. These two competing constraints lead to,

$$\Lambda_{QCD}^2 \ll \mu^2 \ll \left(\frac{\pi}{a} \right)^2 \quad (4.57)$$

- the Rome-Southampton window condition. This applies when we try to calculate and match on a single lattice. Evading this constraint by using multiple lattices and calculating $\sigma_{\mathcal{O}}(\mu, \mu')$ non-perturbatively is our topic for the remainder of this chapter.

4.2 Momentum Sources

For the Rome-Southampton method we need momentum space propagators. These can be calculated explicitly, as in the original Rome-Southampton paper [103], by calculating $S(x, y)$ for sufficiently many point sources and Fourier transforming. However a much more efficient method was introduced in [74] that allows us to obtain the momentum space propagator with a single inversion. Start with the Dirac equation,

$$\sum_z D(x, z) S(z, y) = \delta(x, y) \quad (4.58)$$

where we do the inversion as in section 3.2. We then Fourier transform the y variable,

$$\begin{aligned} \sum_z D(x, z) \sum_y S(z, y) e^{ipy} &= e^{ipx} \equiv \eta(x; p) \\ \sum_z D(x, z) S(z, p) &= \eta(x; p) \end{aligned} \quad (4.59)$$

where,

$$\sum_y S(x, y) e^{ipy} = S(x, p) \quad (4.60)$$

defines $S(x, p)$. If we construct a plane wave source

$$\eta(x; p) = e^{ipx} \delta_{\alpha\beta} \delta_{ab} \quad (4.61)$$

diagonal in spin and colour and the momentum is one of the allowed lattice momenta

$$ap_\mu = n_\mu \frac{2\pi}{L_\mu}, \quad (4.62)$$

with the n_μ integers, solving equation 4.59 gives $S(z, p)$.

To implement the renormalization procedure described in sections 4.1.1 and 4.1.2 for a bilinear operator,

$$O_\Gamma(z) = \bar{q}(z)\Gamma q(z), \quad (4.63)$$

we need the momentum space Green's function $G_\Gamma(p_1, p_2) = S(p_1)\Gamma S(p_2)$ which is,

$$\begin{aligned} G_\Gamma(p_1, p_2) &= \frac{1}{V} \sum_{x,y,z} e^{-ip_1(x-z)} S(x, z) e^{-ip_2(z-y)} \Gamma S(z, y) \\ &= \frac{1}{V} \sum_z S(p_1, z) e^{ip_1 z} \Gamma e^{-ip_2 z} S(z, p_2) \\ &= \frac{1}{V} \sum_z \gamma_5 (e^{-ip_1 z} S(z, p_1))^\dagger \gamma_5 \Gamma e^{-ip_2 z} S(z, p_2), \end{aligned} \quad (4.64)$$

γ_5 Hermiticity is used in the last step. The propagators $S(x, p)$ are found by inverting the Dirac matrix on a plane wave source. We can then multiply by a phase,

$$S'(z, p) = e^{-ipz} S(z, p) \quad (4.65)$$

and form,

$$G_\Gamma(p_1, p_2) = \frac{1}{V} \sum_z \gamma_5 (S'(z, p_1))^\dagger \gamma_5 \Gamma S'(z, p_2). \quad (4.66)$$

Comparing to the first equation in 4.64, this is equivalent to doing the sum over all spatial points, ie. the exact lattice Fourier transform for this momentum. The cost of doing a separate inversion for every momentum is more than offset by the gain in statistical accuracy that comes from doing the volume sum exactly. Once the Green's functions are calculated, the external legs are amputated using the inverse propagator,

$$\Pi_\Gamma(p_1, p_2) = \langle S^{-1}(p_1) \rangle \langle G_\Gamma(p_1, p_2) \rangle (\gamma_5 \langle S^{-\dagger}(p_2) \rangle \gamma_5) \quad (4.67)$$

where

$$S(p) = \frac{1}{V} \sum_x e^{-ipx} S(x, p). \quad (4.68)$$

These are then multiplied by the appropriate projectors, $\mathbf{P}$, to get the bare

vertices,

$$\Lambda_\Gamma = \frac{1}{12} \text{Tr} [\mathbf{P}_\Gamma \Pi_\Gamma]. \quad (4.69)$$

We have written the Green's functions for an arbitrary momentum configuration. For exceptional momenta $p_1 = p_2$ and we have one fewer contraction to do, however, with volume sources only $O(10)$ configurations are needed for extremely high statistical precision, so this is not an important consideration.

4.3 Twisted Boundary Conditions

Momentum sources give very small statistical errors using only a few configurations and sources of systematic error are of more concern. In continuum Euclidean space one has four dimensional rotational symmetry $\mathcal{O}(4)$, the lattice breaks this to $\mathcal{H}(4)$, the symmetry group of a $4d$ hypercube, and we are constrained to use Fourier modes $ap_\mu = n_\mu \frac{2\pi}{L_\mu}$. In order to simulate at different values of $\mu = |p|$ we have to choose different directions which may not be equivalent and the lattice Green's functions have different $O(a^2)$ corrections for different directions. Thus, at non-zero lattice-spacing, Green's functions calculated with $\mathcal{O}(4)$ equivalent but $\mathcal{H}(4)$ inequivalent momenta (eg. $n = (3, 4, 0, 0)$ and $n = (5, 0, 0, 0)$) are different observables. This is a problem, as our intention is to take a continuum limit which is only well defined by taking a single observable across multiple lattices at the same scale. This is difficult or impossible when using Fourier modes.

The $\mathcal{O}(4)$ breaking effects were unimportant when swamped by the statistical error and a smooth interpolation in p^2 could be justified after averaging together $\mathcal{H}(4)$ inequivalent momenta. These effects can be plausibly ignored if we are well below the upper bound of the Rome-Southampton window and $p^2 \ll (\frac{\pi}{a})^2$. In this case anything of $O(a^2)$ is insignificant. In practice this is rarely the case. In order to have a theoretically sound continuum extrapolation where we fix the direction of the momentum and vary the lattice spacing while keeping the physical scale fixed we introduce a new technique - twisted boundary conditions. The Green's functions will still have lattice artefacts but, as long as $\mathcal{H}(4)$ equivalent directions are chosen, they will be the same lattice artefacts and are removed by a continuum extrapolation.

Twisted boundary conditions [30, 17, 44] have been used with success to insert arbitrary momenta into form factor calculations [31, 33]. Here we will use them to

insert arbitrary four momenta and evade the Fourier constraints. Let the quark fields $q(x)$ satisfy twisted boundary conditions [62],

$$q(x + L) = e^{iBx} q(x) \quad (4.70)$$

with,

$$aB_\mu = \frac{\pi\theta_\mu}{L_\mu}. \quad (4.71)$$

and define the twisted quark field,

$$\tilde{q}(x) = e^{-iBx} q(x) \quad (4.72)$$

which satisfies periodic boundary conditions. This changes the continuum Dirac operator,

$$D = (\not{\partial} + m) \rightarrow \tilde{D} = (\not{\partial} + i\not{B} + m). \quad (4.73)$$

Fourier transforming the derivative brings down factors of $p + B$ instead of just p . If the untwisted propagator satisfies the usual Dirac equation, multiplying by the appropriate phase will give the solution to the twisted Dirac equation,

$$\begin{aligned} \sum_z D(x, z) S(z, y) &= \delta(x, y) \\ \rightarrow \sum_z \tilde{D}(x, z) e^{-iB(z-y)} S(z, y) &= \delta(x, y). \end{aligned} \quad (4.74)$$

We will denote the inverse of $\tilde{D}$ by $\tilde{S}$ with

$$e^{-iB(x-y)} S(x, y) = \tilde{S}(x, y). \quad (4.75)$$

We can now introduce momentum sources as before,

$$\sum_z \tilde{D}(x, z) \tilde{S}(z, p) = e^{ipx} \quad (4.76)$$

with

$$\tilde{S}(x, p) = \sum_y \tilde{S}(x, y) e^{ipy} \quad (4.77)$$

where the p' s are Fourier modes. We form the phased propagators,

$$\tilde{S}'(z, p) = e^{-ipz} \tilde{S}(z, p) \quad (4.78)$$

and construct the Green's functions as before,

$$\frac{1}{V} \sum_z \gamma_5(\tilde{S}'(z, p_1))^\dagger \gamma_5 \Gamma \tilde{S}'(z, p_2). \quad (4.79)$$

On substituting we have,

$$\begin{aligned} \tilde{S}'(z, p) &= \sum_y e^{-ip(z-y)} \tilde{S}(z, y) = \sum_y e^{-i(p+B)(z-y)} S(z, y) \\ &= e^{-i(p+B)z} S(z, p+B) = S'(z, p+B) \end{aligned} \quad (4.80)$$

and thus the bilinear Green's function is,

$$\frac{1}{V} \sum_z \gamma_5(S'(z, p_1 + B_1))^\dagger \gamma_5 \Gamma S'(z, p_2 + B_2). \quad (4.81)$$

The crucial thing is that while the p' s still have to be Fourier modes the B' s are completely arbitrary and by changing them appropriately we can reach momenta of any magnitude. For our non-exceptional momentum configuration

$$p_1^2 = p_2^2 = (p_1 - p_2)^2 \quad (4.82)$$

we choose

$$\begin{aligned} n_1 &= (-1, 0, 1, 0) \\ n_2 &= (0, 1, 1, 0) \\ n_1 - n_2 &= (-1, -1, 0, 0). \end{aligned} \quad (4.83)$$

Any lattice momenta related to the choice in 4.83 by a $\mathcal{H}(4)$ symmetry operation is just as good and gives an equivalent observable. We make this choice to minimize $p^{[4]} = \sum_\mu p_\mu^4$ which is a measure of discretization errors and hopefully makes the extrapolation to the continuum limit easier. This comes from the difference between continuum momentum p and lattice momentum $\hat{p}_\mu = \frac{2}{a} \sin\left(\frac{ap_\mu}{2}\right)$; so $p^2 - \hat{p}^2 = \frac{1}{12} a^2 p^{[4]}$ to leading order [16]. It is possible to choose a different direction

and, after a continuum limit is taken, we should reproduce the same results. We have not had the resources to do this but it would be an interesting test.

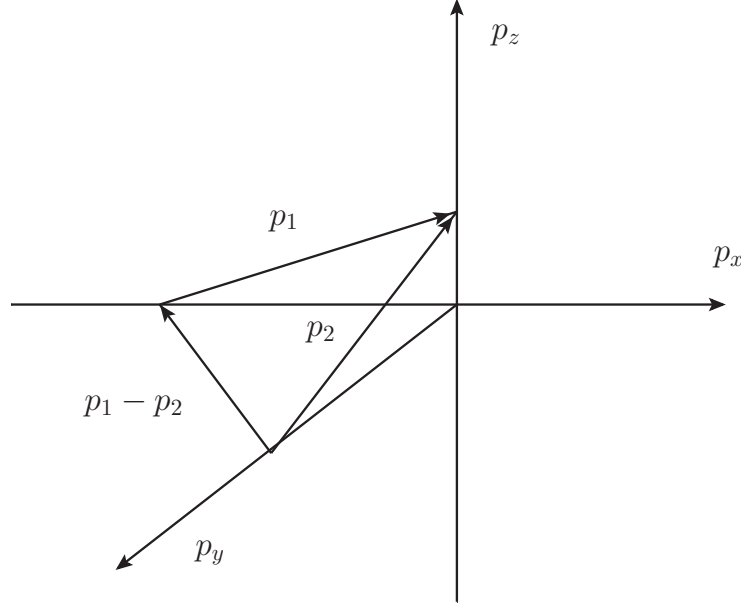


Figure 4.3: Non-exceptional momentum configuration used in this paper. The momenta p_1 , p_2 and $(p_1 - p_2)$ must be equal in magnitude and are represented as an equilateral triangle with vertices touching the p_x , p_y , and p_z axes. With twisted boundary conditions the sides of this triangle can be continuously scaled allowing both smooth interpolation of momentum dependence and controlled continuum extrapolation. This is in contrast to the situation that arises when using Fourier modes.

Figure 4.3 shows that the directions we chose in momentum space form an equilateral triangle. This triangle can be continuously expanded and contracted by varying B . For this particular choice of p we must take

$$\begin{aligned}\theta_1 &= (-\theta, 0, \theta, 0) \\ \theta_2 &= (0, \theta, \theta, 0)\end{aligned}\tag{4.84}$$

and by continuously changing θ we change the magnitude of the momentum while keeping the direction fixed. For exceptional momentum we just choose $p_1 = p_2$ and $B_1 = B_2$. In either exceptional or non-exceptional momentum configurations the vertex function will be a smooth function of $\mu = |p + B|$. It is possible to

simulate directly at any value of momentum or interpolate very accurately and a continuum extrapolation at fixed physical momentum is now possible.

4.4 Four Quark Operators and B_K

We now describe how to renormalize the operator $\mathcal{O}_{VV+AA}$ needed for B_K , as well as other four quark operators, with the Rome-Southampton method [51]. The renormalization condition is the usual one,

$$\frac{Z_{\Gamma\Gamma}}{Z_q^2} \Lambda_{\Gamma\Gamma} = 1 \quad (4.85)$$

where,

$$\{\Gamma\Gamma\} = \{VV \pm AA\}, \{SS \pm PP\}, \{TT\} \quad (4.86)$$

and Λ is the projected amputated vertex function. The renormalization constant for B_K is given by

$$Z_{B_K} = \frac{Z_{VV+AA}}{Z_A^2} = \frac{\Lambda_A^2}{\Lambda_{VV+AA}}. \quad (4.87)$$

The $VV - AA$, $SS \pm PP$ and TT operators mix with $VV + AA$ if chiral symmetry is broken and must be accounted for or shown to be negligible.

The momentum space Green's function for the four quark operators is,

$$G_{\Gamma\Gamma}(p_1, p_2)_{stuv} = \frac{1}{V} \sum_z (\gamma_5(S'(z, p_1))^\dagger \gamma_5 \Gamma S'(z, p_2))_{su} (\gamma_5(S'(z, p_1))^\dagger \gamma_5 \Gamma S'(z, p_2))_{tv} \quad (4.88)$$

with the primes denoting the phased propagators of equation 4.78 and the indices $\{s, t, u, v\}$ representing the twelve spin and colour components of each propagator. The amputated four point vertex is,

$$\Pi_{\Gamma\Gamma}(p_1, p_2)_{abcd} = (\gamma_5(S^{-1})^\dagger(p_2) \gamma_5)_{as} (\gamma_5(S^{-1})^\dagger(p_2) \gamma_5)_{bt} G_{\Gamma\Gamma}(p_1, p_2)_{stuv} S_{uc}^{-1}(p_1) S_{vd}^{-1}(p_1) \quad (4.89)$$

with,

$$S_{ab}(p) = \frac{1}{V} \sum_x e^{-ipx} S'_{ab}(x, p). \quad (4.90)$$

The projectors onto the tree level values are given by,

$$\mathbf{P}_{(\gamma_\mu),\alpha\beta\gamma\delta}^{ijkl} = \frac{1}{256} [(\gamma^\nu)_{\beta\alpha}(\gamma_\nu)_{\delta\gamma} + (\gamma^\nu\gamma_5)_{\beta\alpha}(\gamma_\nu\gamma_5)_{\delta\gamma}] \delta_{ij} \delta_{kl} \quad (4.91)$$

and

$$\mathbf{P}_{(\not{q}),\alpha\beta\gamma\delta}^{ijkl} = \frac{1}{64q^2} [(\not{q})_{\beta\alpha}(\not{q})_{\delta\gamma} + (\not{q}\gamma_5)_{\beta\alpha}(\not{q}\gamma_5)_{\delta\gamma}] \delta_{ij} \delta_{kl}, \quad (4.92)$$

where $ijkl$ are colour and $\alpha\beta\gamma\delta$ are spin indices. They define two different renormalization schemes for the four quark operators, similarly to RI and RI'. Then the (projected and amputated) vertex functions are,

$$\Lambda_{\Gamma\Gamma}^{(X)} = \frac{1}{12} \text{Tr} [\mathbf{P}_{(X),\alpha\beta\gamma\delta}^{ijkl} \Pi_{\Gamma\Gamma,\alpha\beta\gamma\delta}^{ijkl}] \quad (4.93)$$

where $X = \{\not{q}, \gamma_\mu\}$. This defines the vertex function in the two schemes.

To fully specify the renormalization condition we must choose the condition on Z_q . We have two candidates: defining Z_q through the vector vertex function,

$$Z_q^{(\not{q})} = \frac{1}{12} \text{Tr} [\frac{q^\mu}{q^2} \not{q} \Pi_V^\mu]; \quad (4.94)$$

or

$$Z_q^{(\gamma_\mu)} = \frac{1}{12} \text{Tr} [\frac{1}{4} \gamma^\mu \Pi_V^\mu] \quad (4.95)$$

corresponding to the definitions in the RI-SMOM and RI-SMOM $_{\gamma_\mu}$ schemes respectively. Thus there are four possibilities,

$$Z_{\Gamma\Gamma}^{(\gamma_\mu, \not{q})} = (\Lambda_{\Gamma\Gamma}^{(\gamma_\mu)})^{-1} (Z_q^{(\not{q})})^2 \quad (4.96)$$

$$Z_{\Gamma\Gamma}^{(\not{q}, \not{q})} = (\Lambda_{\Gamma\Gamma}^{(\not{q})})^{-1} (Z_q^{(\not{q})})^2 \quad (4.97)$$

$$Z_{\Gamma\Gamma}^{(\gamma_\mu, \gamma_\mu)} = (\Lambda_{\Gamma\Gamma}^{(\gamma_\mu)})^{-1} (Z_q^{(\gamma_\mu)})^2 \quad (4.98)$$

$$Z_{\Gamma\Gamma}^{(\not{q}, \gamma_\mu)} = (\Lambda_{\Gamma\Gamma}^{(\not{q})})^{-1} (Z_q^{(\gamma_\mu)})^2 \quad (4.99)$$

and we have no reason to prefer any particular one so we will use all four and pick the one where the non-perturbative running agrees most closely with perturbation theory as the best. The exceptional scheme uses,

$$Z_{\Gamma\Gamma}^{(RI-MOM)} = (\Lambda_{\Gamma\Gamma}^{(\gamma_\mu)})^{-1} (Z_q^{(\gamma_\mu)})^2 \quad (4.100)$$

where the vertex functions on the right are evaluated at the exceptional kinematic point. This is expected to show larger chiral symmetry breaking effects and the mixing of $VV + AA$ with the other operators will be more significant.

4.5 Step Scaling

Matching with perturbation theory in the Rome-Southampton window is difficult due to the low energy scales at which we are simulating and the large coupling this implies. It would be better to raise the scale non-perturbatively and convert to $\overline{MS}$ only at the high scale, where the coupling is smaller and perturbative convergence is better. To this end we define the ratios,

$$R_{\mathcal{O}}(\mu, a, m) = \frac{\Lambda_A(\mu, a, m)}{\Lambda_{\mathcal{O}}(\mu, a, m)} = \frac{Z_{\mathcal{O}}(\mu, a, m)}{Z_A} \quad (4.101)$$

which are extrapolated to the chiral limit to give the renormalization constant,

$$Z_{\mathcal{O}}(\mu, a) = Z_A \lim_{m \rightarrow 0} R_{\mathcal{O}}(\mu, a, m). \quad (4.102)$$

We have used the local axial current vertex to eliminate factors of Z_q in favour of factors of Z_A which we compute from the domain wall conserved current, section 1.4.1. Then the continuum limit of the renormalized operator can be taken,

$$\langle \mathcal{O} \rangle(\mu) = \lim_{a \rightarrow 0} \langle \mathcal{O}(a) \rangle Z_{\mathcal{O}}(\mu, a). \quad (4.103)$$

The scale μ , set by the momentum, must be the same across all lattices used to take the continuum limit. All of the scale dependence is carried by the renormalization constant. Thus, to convert $\langle \mathcal{O} \rangle(\mu)$ to $\langle \mathcal{O} \rangle(s\mu)$, where the scale factor, s , is a real number, we multiply by,

$$\Sigma_{\mathcal{O}}(\mu, s\mu, a) = \lim_{m \rightarrow 0} \frac{R_{\mathcal{O}}(s\mu, a, m)}{R_{\mathcal{O}}(\mu, a, m)}, \quad (4.104)$$

which has,

$$\sigma_{\mathcal{O}}(\mu, s\mu) = \lim_{a \rightarrow 0} \Sigma_{\mathcal{O}}(\mu, s\mu, a) \quad (4.105)$$

as its continuum limit.

$\sigma_{\mathcal{O}}(\mu, s\mu)$ tells us the scale dependence, in a fixed scheme, in the continuum

limit around the scale μ and we have given an explicit way to construct the ratios of section 4.1.3 non-perturbatively. Now, using continuum limit, non-perturbative running given by the σ 's, we can assess how well perturbation theory describes the running at low energy scales. The twisted boundary conditions allow us to pick a fixed momentum orientation for the amplitude as a function of a , so there is no ambiguity in taking the continuum limit from $\mathcal{H}(4)$ breaking effects. We will calculate these functions later in the various regularization invariant schemes.

4.5.1 Previous Work on Step Scaling

Step scaling is the natural approach to connecting low and high energies in lattice QCD. The brute force simulation of all scales on a single lattice is avoided by having a series of simulations where the physical volume is reduced with the lattice spacing. This is well established in the Schrödinger functional scheme [100, 102, 101, 99, 124, 123, 25] and there is also a scheme based on Creutz ratios aiming to calculate the non-perturbative running of α [21]. In our approach, in contrast to the Schrödinger functional, the renormalization scale is not tied to the volume. This means we do not need to tune the bare couplings to precisely match volumes between simulations.

Step scaling with RI/MOM has been previously studied [139] (see also [50]). Reference [139] used a series of quenched configurations where the ratio of lattice spacings had been tuned to be exactly two in order to align the Fourier modes. In a dynamical simulation this fine tuning would be expensive. Power counting and the Rome-Southampton scaling window were used to justify neglecting lattice artefacts rather than the controlled continuum extrapolation at fixed physical momenta, as in this thesis. Free parameters for each operator were introduced to match the renormalisation constants for different β 's corresponding to the (possibly non-perturbative) anomalous scaling of the operator with the lattice cut-off. This excellent start was not easily developed into a practical technique and improvements in RI-MOM technology now enable us to do a better job.

Schrödinger Functional

We describe the step scaling method of [102] that has been used very successfully, together with the Schrödinger functional (SF) [101], to calculate the running of the gauge coupling non-perturbatively. For fixed bare parameters, in the massless

case this just means fixed β , the SF coupling $g(\mu = \frac{1}{aL})$ is measured on an L^4 lattice and $g(\mu' = \frac{1}{saL})$ is measured on an $(sL)^4$ lattice, where s is a scale factor that typically equals 2, L is an integer and aL is the box length in physical units. This gives,

$$g(\mu', a) = \Sigma(s, g(\mu), \frac{a}{L}), \quad (4.106)$$

that is, at fixed β (which is equivalent to fixed lattice spacing), Σ tells us how $g(\mu)$ changes due to a scale change by a factor s . The continuum limit is taken holding $g(\mu)$ fixed, which means changing L and tuning β (and hence a) to match the volumes and give the correct value of $g(\mu)$. The continuum limit gives the step scaling function

$$\sigma(s, g(\frac{1}{L})) = \lim_{a \rightarrow 0} \Sigma(s, g(\mu), \frac{a}{L}) = \lim_{a \rightarrow 0} g(\mu' = \frac{1}{saL}, a). \quad (4.107)$$

The coupling $g(\mu)$ is the physical quantity that is fixed as we take the continuum limit and defines the physical volume. The SF allows it to be defined at the scale $\mu = \frac{1}{La}$, the most infra-red scale accessible on a single lattice.

The first step evolves the coupling from $\mu'_1 \rightarrow \mu_1$, the next step needs to evolve it from $\mu'_2 \rightarrow \mu_2$ with $\mu_2 = \mu'_1$. In the second step the quantity $g(\mu_2)$ is held fixed for the continuum limit compared to $g(\mu_1)$ in the first and consecutive steps will only join up in the continuum limit if $g(\mu_2, a) = \lim_{a \rightarrow 0} g(\mu'_1, a)$. The quantity used to determine the continuum limit changes at each step so that it is appropriate to the energy scales considered. The SF coupling, g , is directly coupled to the volume and matching it requires fine tuning β for many different values of L . Our step scaling scheme avoids this fine tuning of β but the most infra-red scale accessible to us in a single step will be larger than $\frac{1}{La}$ to satisfy the step scaling window condition, see section 4.5.2. Thus we avoid tuning β at the expense of shorter steps.

4.5.2 RI-MOM Step Scaling Method

We have continuum limit non-perturbative scale evolution ratios that can be used to access higher momentum scales, the functions $\sigma_{\mathcal{O}}(\mu, s\mu)$. Step scaling combines a sequence of these objects, each determined inexpensively on small lattices, to raise the renormalization scale from μ to $s\mu$. Three technical advantages we have over the earlier attempt at RI/MOM step scaling [139] make the task

more tractable. First, volume sources greatly reduce statistical errors. Second, non-exceptional momenta render mass dependence and infra-red behaviour in p^2 benign; this greatly assists with both chiral extrapolation and matching between different lattices and, in addition, reduces finite volume sensitivity. Third, twisted boundary conditions are used to select arbitrary four momenta. The direction of the momenta is kept fixed relative to lattice axes and this allows the same physical momenta be chosen on each ensemble with the same lattice artefacts. The vertex functions then have a valid expansion in powers of a , enabling precise matching of momenta between ensembles and precise continuum extrapolation.

On each lattice we will keep the off-shell momentum scales hard enough that the physical volume should not be resolved. In this way the calculation will be finite volume effect safe and the perturbative matching will be done with standard, infinite volume perturbation theory, already available to high order for RI schemes. The use of infinite volume perturbation theory, the fact we don't need to tune the bare coupling and the ease with which we can extend the technique to general operators are significant advantages over the Schrödinger functional step scaling schemes.

For the step scaling to be affordable we need to be able to simulate in a small physical volume. For vertex functions whenever

$$\mu^2 \gg \left(\frac{\pi}{La}\right)^2 \quad (4.108)$$

the process takes place at distances too small to resolve the finite dimensions of the box. We also have the condition that, on a single lattice, the momentum must be smaller than the cutoff, so that the lattice artefacts are small. This gives us the “step scaling window”

$$\left(\frac{\pi}{La}\right)^2 \ll \mu^2 \ll \left(\frac{\pi}{a}\right)^2 \quad (4.109)$$

in place of the Rome Southampton window. Until we have taken a continuum limit and raised the renormalization scale to a regime where the coupling is small perturbation theory is forgotten. The step scaling window condition should be possible to satisfy with modest lattices, something of the order $L \leq 16$ at all stages of the calculation.

We will compute vertex functions on lattices at different values of β with

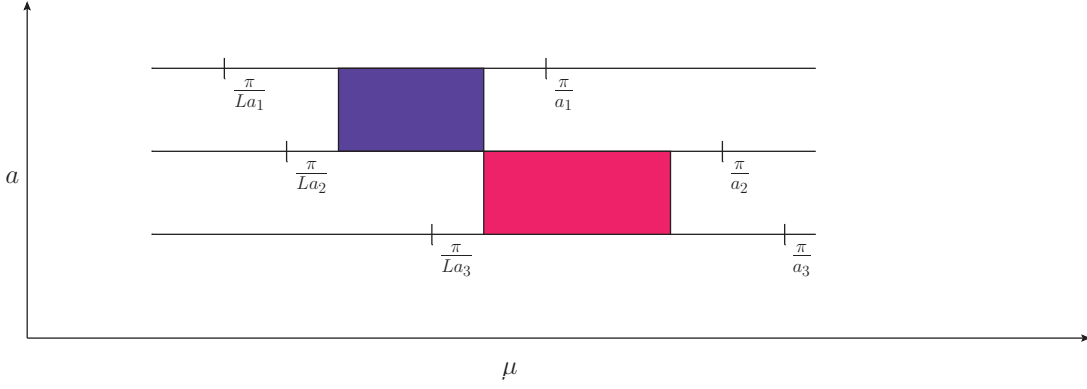


Figure 4.4: Schematic picture of the step scaling process. Each long horizontal line signifies a different lattice, with lattice spacing decreasing top to bottom. Starting at the top left we have the step scaling window (the region of μ within the vertical ticks) for the first lattice $\frac{\pi}{La_1} \ll \mu \ll \frac{\pi}{a_1}$ which overlaps with the window of the second lattice in the purple region, we obtain $\sigma(\mu, s\mu)$ from continuum extrapolation between these lattices where μ is the left of the purple window and $s\mu$ is the right. We do the same between lattices two and three and the pink window starting at $s\mu$ and finishing at $s^2\mu$. Note that since we have continuous tuning of our scale we could finish at $s_1 s_2 \mu$ where s_1, s_2 are different.

overlapping scaling windows, see figure 4.4 for a picture of this. Between each pair of lattices we will calculate the step scaling function $\sigma_{\mathcal{O}}(\mu, s\mu)$ for some operator $\mathcal{O}$ between scales μ and $s\mu$ where both are chosen inside the step scaling window. These are then combined in a series of non-perturbative steps from the low scale μ to the high one $s^n\mu$. The perturbative conversion to $\overline{MS}$ is then applied only where it is well convergent at $\mu' = s^n\mu$,

$$\begin{aligned}
 \langle \mathcal{O}^{\overline{MS}}(\mu') \rangle &= \langle \mathcal{O}^{\text{SMOM}}(\mu) \rangle \\
 &\times \sigma_{\mathcal{O}}^1(\mu, s\mu) \\
 &\times \sigma_{\mathcal{O}}^2(s\mu, s^2\mu) \\
 &\dots \\
 &\times \sigma_{\mathcal{O}}^n(s^{n-1}\mu, s^n\mu) \\
 &\times C^{\text{SMOM} \rightarrow \overline{MS}}(\alpha_s(\mu' = s^n\mu)).
 \end{aligned} \tag{4.110}$$

There appear to be no obstacles to satisfying the step scaling condition at reasonable expense. For example $L^3 = 16^3$ domain wall fermion simulations will be inexpensive with the next generation of supercomputers; multiple ensembles

dedicated solely to the renormalisation of lattice operators are quite affordable. This is particularly helped by the relatively benign mass dependence of non-exceptional momentum vertex functions, so that only a few masses need be simulated and there is no need to for very small mass dynamical fermions. It is also clear that on these proposed small ensembles all hadronic quantities must be avoided since they will suffer greatly from finite volume artefacts, the boxes will be quite a bit smaller than a hadron. For the determination of m_{res} , PCAC is an operator relation and so it is immaterial whether a finite volume or a physical pionic state is used to determine the ratio. Determining the lattice spacing (or more specifically the ratio of the lattice spacing to that of an earlier larger volume simulation) in a finite volume safe manner is harder, a proposal based on the inter-quark potential is discussed below.

If the lattice spacing could be measured sufficiently accurately at all length scales there would be no barrier to doing three flavour step scaling up to any energy. However it is our intention to match the lowest step in the sequence with a large volume, low energy lattice calculation and then use the step scaling functions to convert quantities renormalized at low scales to high scales. At high enough energy we would have to have four and five flavour running as we cross the charm and bottom thresholds. We do not yet have access to lattices with dynamical charm and so restrict our attention to three flavour step scaling in a regime where the effects of charm quarks should be small $< 5\text{GeV}$.

Continuum Limit

We described in section 4.5.1 how, using the Schrödinger functional, a different physical quantity at each step can be used to determine the physical length scale. In a similar way we seek a family of quantities, $\{q_n\}$, that can be used to determine the lattice spacing at each step, analogous to the $g(\frac{\mu}{s^n})$ in the Schrödinger functional. In just the same way as with $g(\frac{\mu}{s^n})$ in the Schrödinger functional, the continuum limit of each q_n will be determined in order to make sure successive steps join up. At step n we will determine the continuum limit of q_{n+1} while determining the lattice spacing $a^{(q_n)}$ from q_n via

$$q_n(a) = q_n^{\text{cont}}, \quad (4.111)$$

where q_n^{cont} was determined in the previous step. This is most simply seen if q has dimensions of energy. The lower scale at step n , μ_n , measured in units of q_n must, in the continuum limit, equal the upper scale from step $n - 1$ measured in the same units,

$$\frac{\mu_n}{q_n(a)} = \frac{\mu_n}{q_{n-1}^{\text{cont}}} \left(\frac{q_{n-1}}{q_n} \right)^{\text{cont}} \quad (4.112)$$

which is true if equation 4.111 is used. At step n , q_{n+1}^{cont} is found from,

$$q_{n+1}^{\text{cont}} = \lim_{a(q_n) \rightarrow 0} q_{n+1}(a^{(q_n)}) \Big|_{q_n(a)=q_n^{\text{cont}}} . \quad (4.113)$$

The scale evolution functions σ^n are,

$$\sigma_{\mathcal{O}}^n(s^{n-1}\mu, s^n\mu) = \lim_{a(q_n) \rightarrow 0} \Sigma_{\mathcal{O}}^n(s^{n-1}\mu, s^n\mu, a^{(q_n)}) \Big|_{q_n(a)=q_n^{\text{cont}}} . \quad (4.114)$$

We want q to be a continuously variable scale so we can change it, post simulation, to always be finite volume safe and avoid fine tuning β . Eventually, for small enough volume, the dependence of q on the lattice spacing will become logarithmic and we will lose precision, but several steps should be possible before this occurs.

In principle we can use the momentum dependence of the off-shell vertex functions themselves to match the lattice scales between different ensembles. This may be assisted by the improved techniques of this thesis, particularly the tunable momentum scale. This allows us to select the length scale that sets lattice spacing post simulation. However, initial explorations of this idea have not been very promising. If we accept fine tuning β it is also possible to combine step scaling of off-shell vertex functions with existing finite volume schemes, such as the Schrödinger functional or the Wilson loop scheme. In this thesis we chose the static potential and the Sommer scale as the best physical quantities to use that do not depend on hadronic physics and can be measured sufficiently accurately. Similar tunable scales, in particular one based on the Wilson flow [28], also appear promising, however these have not been investigated in this thesis.

Static Potential

The static potential and the Sommer parameter, see section 3.5, have been successfully measured over a range of length scales for the Wilson action [106].

In this reference the Sommer parameter $r_0 \approx 0.48 fm$ was replaced with a smaller distance parameter $r \approx 0.26 fm$. This gave a small distance quantity with which to set the scale, for our purposes as we go to smaller volumes we will need a quantity unaffected by long distance scales, if our box size is smaller than $0.5 fm$ then r_0 is not a good unit to measure distances. A whole sequence of scales similar to the Sommer scale can be defined using,

$$r_n^2 F(r_n) = C_n. \quad (4.115)$$

The Sommer scale r_0 takes

$$C_0 = 1.65 \quad (4.116)$$

which we can use for our largest volume simulations. This provides an explicit example of the physical quantities $\{q_n\}$ from the previous section. A step scaling scheme with scale factor s can then be defined choosing $\mu_n = s^n \mu$ and $r_n = \frac{r_0}{s^n}$ as follows:

- Determine $\sigma(\mu_0, \mu_1)$ in continuum limit holding $r_0 \mu_0$ fixed such that $r_0^2 F(r_0) = C_0$
- Determine $C_1 = \frac{r_0^2}{s^2} F(\frac{r_0}{s})$ in continuum limit holding r_0 fixed
- Decrease the box size by $\simeq \frac{1}{s}$ without fine tuning
- Determine $\sigma(\mu_1, \mu_2)$ in continuum limit holding $r_1 \mu_1$ fixed such that $r_1^2 F(r_1) = C_1$
- Determine $C_2 = \frac{r_1^2}{s^2} F(\frac{r_1}{s})$ in continuum limit holding r_1 fixed
- Decrease the box size by $\simeq \frac{1}{s}$ without fine tuning
- etc...

For convenience we denote the scale factor at every step by s , it is of course possible to change it from step to step if this is necessary. Following the rule of thumb that $r < \frac{L}{3}$ for the static potential should ensure finite volume safety and enable simulation down to $L \simeq 0.75 fm$. Eventually this will become imprecise when entering a region where the potential runs logarithmically. However, having step scaling functions down to this distance would already be of substantial benefit.

4.6 Results on Small Lattices

We now calculate momentum source Green's functions with twisted boundary conditions on the $16^3 \times 32$, $L_s = 16$, $M_5 = 1.8$, Iwasaki gauge $\beta = 2.13$ configurations with $m_s = 0.04$ described in [1]. There are three ensembles with degenerate light quark masses, $m = \{0.01, 0.02, 0.03\}$, we calculate on each and make a linear extrapolation to the chiral limit $m = -m_{res} = -0.00305$ and use $Z_A = 0.7161(1)$, $a^{-1} = 1.729(1) \text{ GeV}$ [2]. A second ensemble set with $\beta = 2.23$ was used which has not previously been published. This was part of a parameter search, made in the style of [6, 5], prior to the 32^3 simulations in [10] which lie nearby in parameter space. There are two ensembles with $m_s = 0.04$ and $m = \{0.01, 0.02\}$ with around 2000 trajectories in each ensemble. We estimate $m_{res} \simeq 10^{-3}$ and $Z_A = 0.740$ by interpolating between nearby values for β (specifically those from [10]) which is not ideal but suffices for an exploratory calculation. For step scaling functions we do not need Z_A and we will see that the mass dependence is very weak so a relatively large error in m_{res} is acceptable.

4.6.1 Vertex Functions

All the data in this section was computed using the volume source method with twisted boundary conditions to reach the appropriate momenta, because of the statistical accuracy only 20 configurations on each ensemble were needed, see tables 4.1 and 4.2 with momenta in table 4.3.

am_q	Range	Δ
0.03	800 - 2320	80
0.02	1000 - 2520	80
0.01	1000 - 2520	80

Table 4.1: $\beta = 2.13$ lattice giving the range of molecular dynamics time and the separation Δ between gauge configurations used in this work. Because so few configurations are needed with the volume averaging a large Δ was chosen to minimise auto-correlation effects.

First we make a comparison between the vertex functions Λ_O computed with and without twisted boundary conditions. We plot the vertex functions for Λ_A and Λ_V in figure 4.5. They have been calculated in two ways: choosing the

am_q	Range	Δ
0.02	1000 - 1760	40
0.01	1240 - 2000	40

Table 4.2: $\beta = 2.23$ lattice giving the range of molecular dynamics time and the separation Δ between gauge configurations used in this work.

β	ap_1	ap_2	Range
2.13	$(0, x, x, 0)$	$(-x, 0, x, 0)$	$x = \frac{2\pi}{L}(1.875) \rightarrow x = \frac{2\pi}{L}(2.75)$
2.23	$(0, x, x, 0)$	$(-x, 0, x, 0)$	$x = \frac{2\pi}{L}(1.5625) \rightarrow x = \frac{2\pi}{L}(3.125)$

Table 4.3: The momenta at which we calculated the vertex functions. We note that the momentum $\frac{2\pi}{L}(0, 1.875, 1.875, 0)$, for example, can be reached using ‘base’ momentum $\frac{2\pi}{L}(0, 1, 1, 0)$ and twist $\frac{2\pi}{L}(0, 0.875, 0.875, 0)$ or base $\frac{2\pi}{L}(0, 2, 2, 0)$ and twist $\frac{2\pi}{L}(0, -0.125, -0.125, 0)$. We chose the base momentum such that the magnitude of the twisting component was less than $\frac{2\pi}{L}(0.75)$, but such a choice is entirely arbitrary.

momenta using different Fourier modes chosen to minimize $p^{[4]} = \sum_i p_i^4$, which breaks $\mathcal{H}(4)$ symmetry; using twisted boundary conditions to fix the direction and vary the magnitude only, which preserves $\mathcal{H}(4)$. The scatter in the data caused by $\mathcal{H}(4)$ breaking in the first method does not occur when using twisted boundary conditions, as expected. Lattice artefacts are still present and of $O(a^2)$, but they will now vanish as a consistent continuum limit is taken of the same observable across multiple lattices. Note also that although we have used very few configurations, we still have very good statistical accuracy, demonstrating the power of the momentum source technique.

We now show the vertex functions for all of the bilinears as well as the operator O_{VV+AA} relevant for B_K . The projectors are defined in 4.1.1, 4.1.2 and we use exceptional and non-exceptional schemes. We will always use RI-MOM for exceptional kinematics and not RI'-MOM, the results at exceptional momenta will not be used for any final values in this thesis because the non-exceptional schemes have much smaller systematic errors. In figure 4.6 we show the chiral limit for the axial vector vertex function for several values of momentum. The data is quite flat and we do not see any significant curvature, especially in the non-exceptional case at high momentum. We take this as an indication that our momenta are high enough to suppress poles in the quark mass and that we can

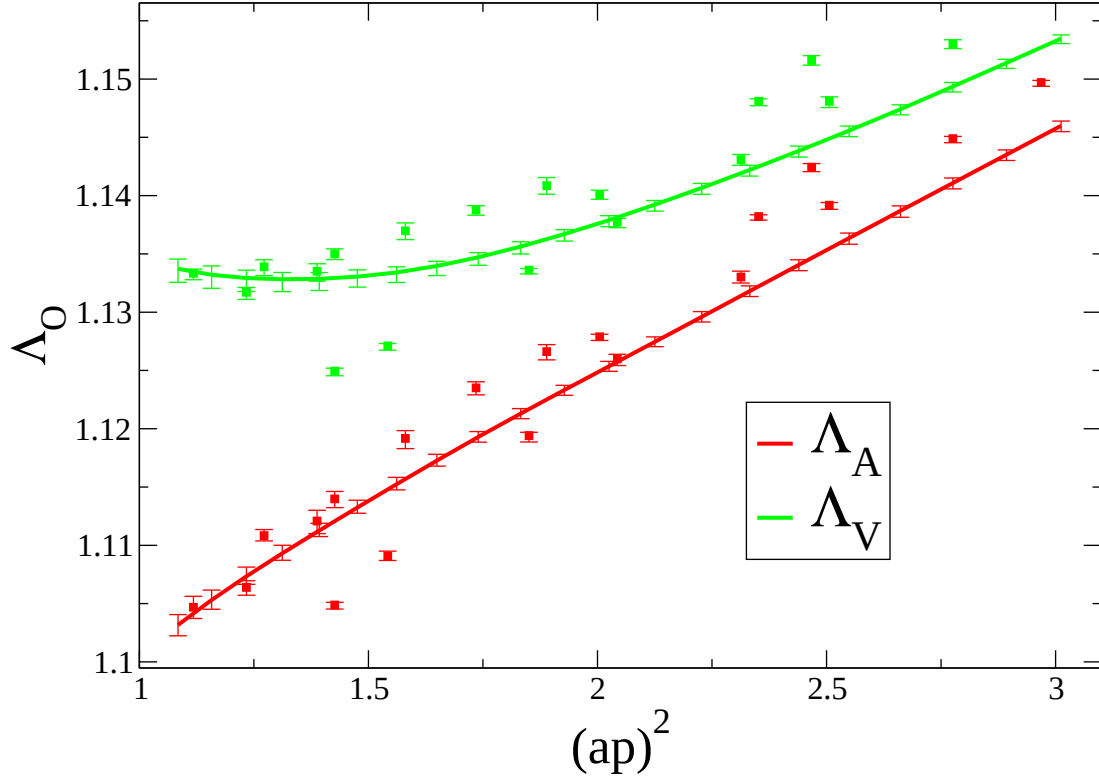


Figure 4.5: The axial (red) and vector (green) exceptional vertex functions on the $\beta = 2.13$ lattice at $m_q = 0.03$. Comparing untwisted (squares) with twisted (connected) data; twisting completely eliminates the $\mathcal{H}(4)$ breaking scatter.

linearly extrapolate in the mass.

Figure 4.7 shows the chiral limit of the bilinear vertex functions and the vertex Λ_{VV+AA} using the projectors with and without a $\not{q}$ in the non-exceptional case. Note that in the non-exceptional schemes we see near equality between Λ_A and Λ_V , figure 4.8, even at small values of momentum, as expected from the extra suppression of spontaneous symmetry breaking effects in this scheme. The vertex functions for the scalar and pseudoscalar operators are also much closer in the non-exceptional schemes.

Our renormalization condition is

$$\frac{Z_\sigma}{Z_q} \Lambda_\sigma = 1, \quad (4.117)$$

we drop the B subscript of section 4.1.1, all vertex functions will be bare vertex

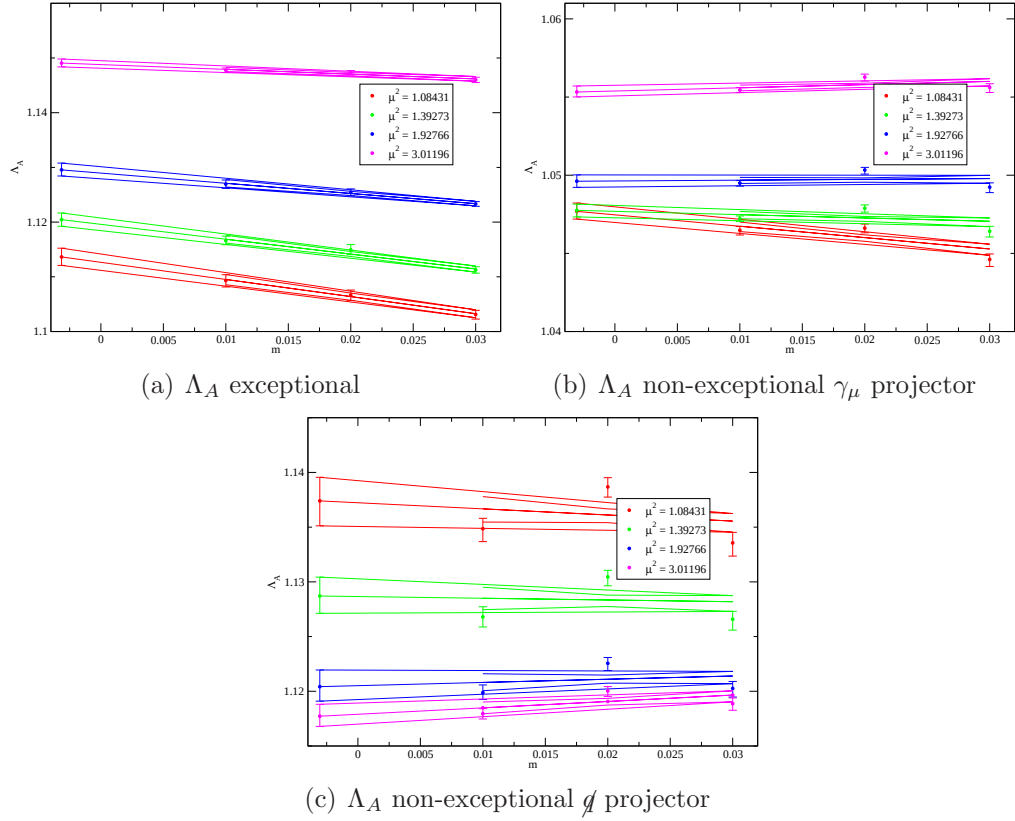


Figure 4.6: The chiral extrapolation of the axial vertex function on the $\beta = 2.13$ lattice in three schemes. The chiral limit seems very benign especially at larger and non-exceptional momentum. A similar picture is observed for the other operators $\{V, S, P, T, VV + AA\}$.

functions from here onwards. From $\Lambda_{\mathcal{O}}$ we can calculate the renormalization constants in the various RI schemes. We follow the strategy of [8] and eliminate factors of Z_q in favour of factors of Z_A using

$$\frac{Z_q}{Z_A} = \Lambda_A. \quad (4.118)$$

and use the known value of Z_A . The Ward identities tell us that $\Lambda_A = \Lambda_V$ and $\Lambda_S = \Lambda_P$ which, as we have seen, holds in the non-exceptional case but not so well in the exceptional case. The renormalization constants that we can compute

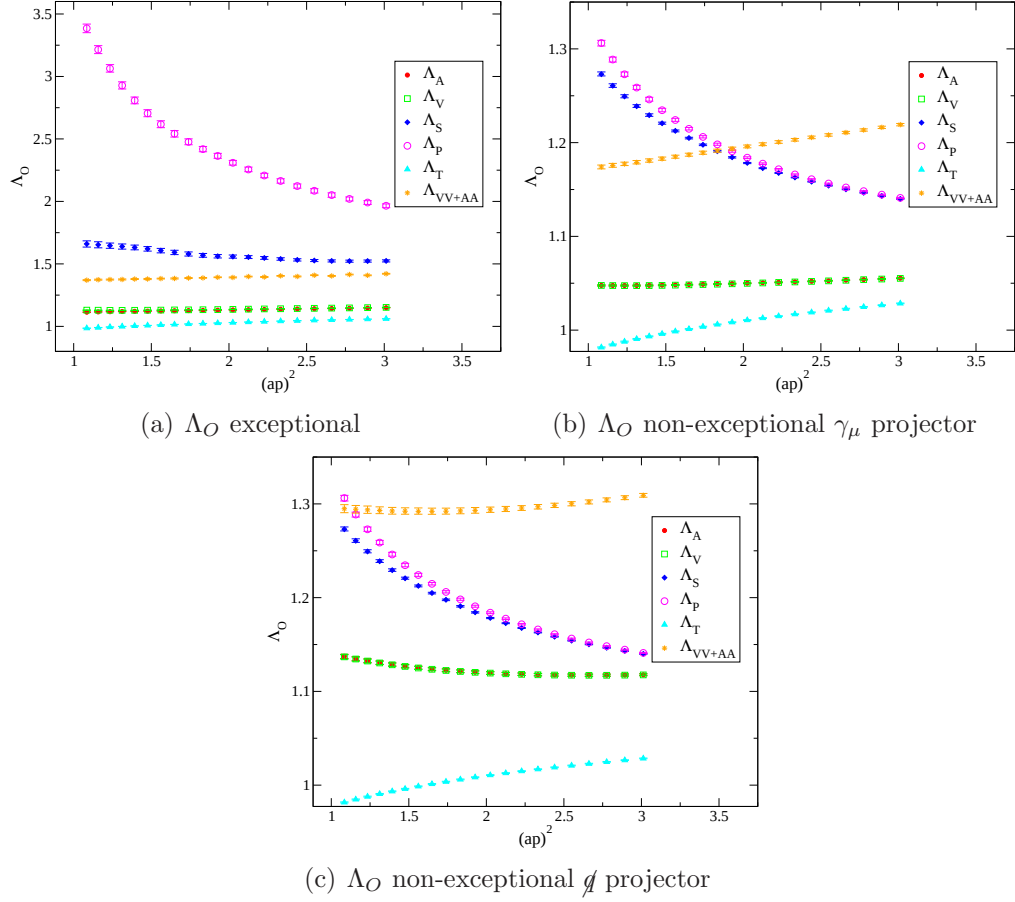


Figure 4.7: All the bilinear vertex functions and the $VV+AA$ four quark operator calculated on the $\beta = 2.13$ lattice in the chiral limit. Note the great reduction in chiral symmetry breaking effects ($\Lambda_A - \Lambda_V$ and $\Lambda_S - \Lambda_P$) at non-exceptional momentum.

from the bilinears and Λ_{VV+AA} are Z_q as above and

$$Z_m = \frac{\Lambda_S}{Z_A \Lambda_A} \quad (4.119)$$

$$Z_T = Z_A \frac{\Lambda_A}{\Lambda_T} \quad (4.120)$$

$$Z_{BK} = \frac{Z_{VV+AA}}{Z_A^2} = \frac{\Lambda_A^2}{\Lambda_{VV+AA}} \quad (4.121)$$

in the RI-SMOM and RI-SMOM $_{\gamma_\mu}$ schemes. In the RI-MOM scheme we use $\frac{1}{2}(\Lambda_A + \Lambda_V)$ in place of Λ_A for the central value to match old references more closely, ie. [8]. The renormalization constants in each of the schemes are the red

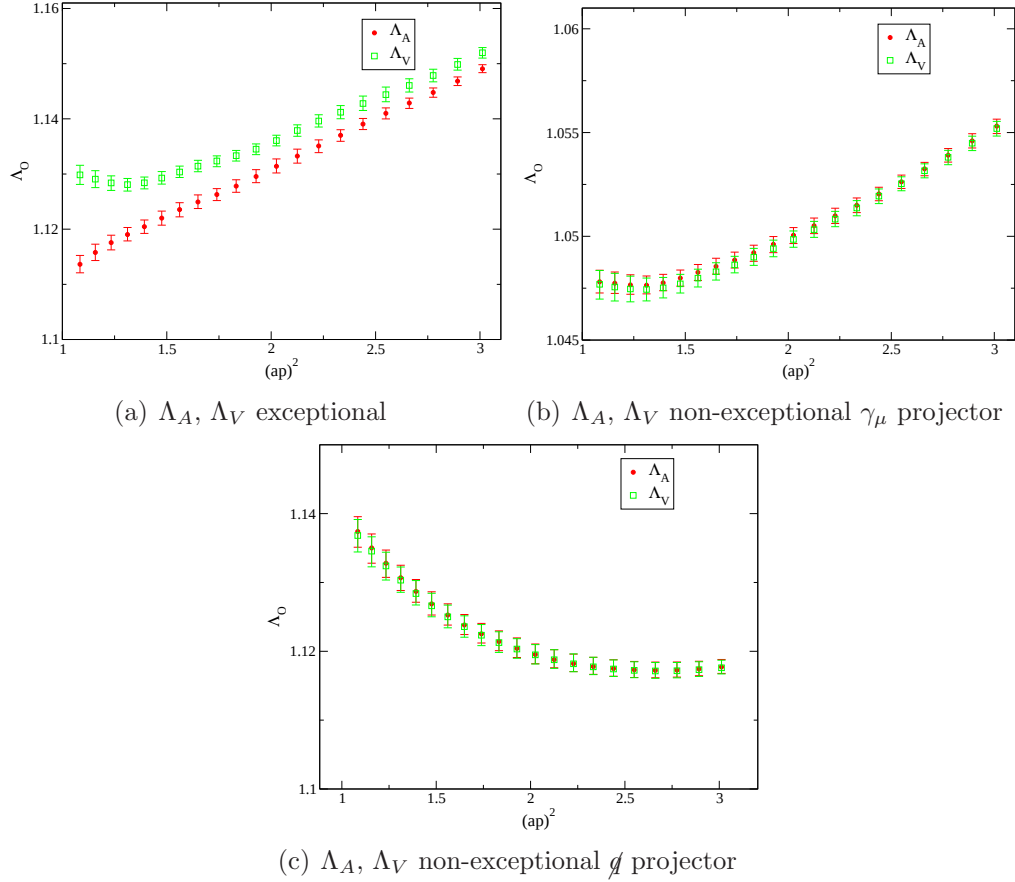


Figure 4.8: Zoomed in view of Λ_A and Λ_V in the chiral limit with non-exceptional and exceptional kinematics from figure 4.7. Note the finer scale in the non-exceptional plot. The roughly 2% difference in the exceptional case is eliminated, even at low scales, for non-exceptional kinematics.

data in figures 4.9, 4.10, 4.11 and 4.12.

We can use the perturbative running to convert the RI results to $\overline{MS}$ as in section 4.1.3. We use the perturbative result, cf. equation 4.56,

$$\frac{Z_O^{(S)}(\mu')}{Z_O^{(S)}(\mu)} = \exp \left(\int_{\alpha(\mu)}^{\alpha(\mu')} \frac{\gamma^{(S)}(x)}{\beta^{(S)}(x)} dx \right). \quad (4.122)$$

where S is the renormalization scheme. To convert from

$$S = \{\text{RI-MOM}, \text{RI-SMOM}, \text{RI-SMOM}_{\gamma_\mu}\}$$

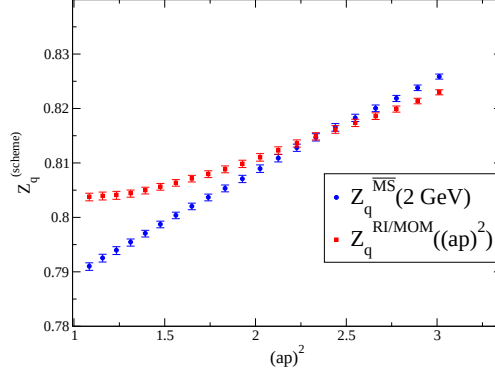
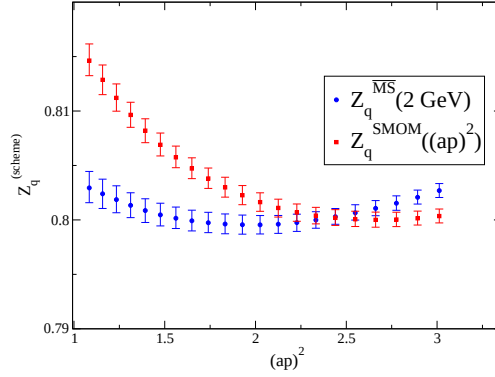
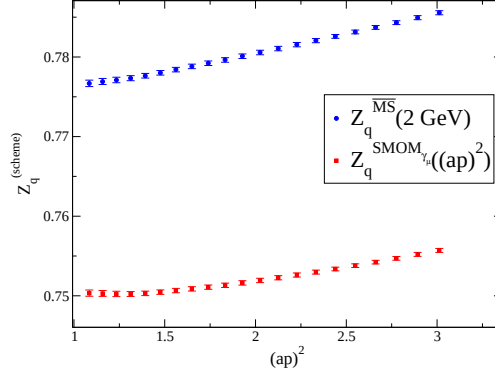
(a) $Z_q^{RI/MOM}((ap)^2) \rightarrow Z_q^{\overline{MS}}(2\text{GeV})$ (b) $Z_q^{SMOM-\not{g}}((ap)^2) \rightarrow Z_q^{\overline{MS}}(2\text{GeV})$ (c) $Z_q^{SMOM-\gamma_\mu}((ap)^2) \rightarrow Z_q^{\overline{MS}}(2\text{GeV})$

Figure 4.9: Z_q on the $\beta = 2.13$ lattice in a MOM scheme (red) and with the perturbative running divided out and converted to $\overline{MS}$ (blue). Z_q contains strong lattice artefacts for domain wall fermions and runs in the opposite direction to the perturbative prediction at this lattice spacing. We find later that this changes in the continuum limit.

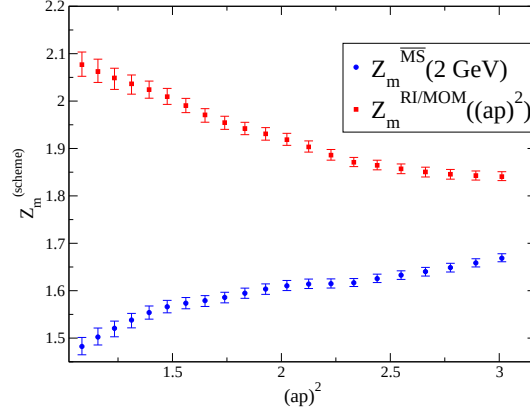
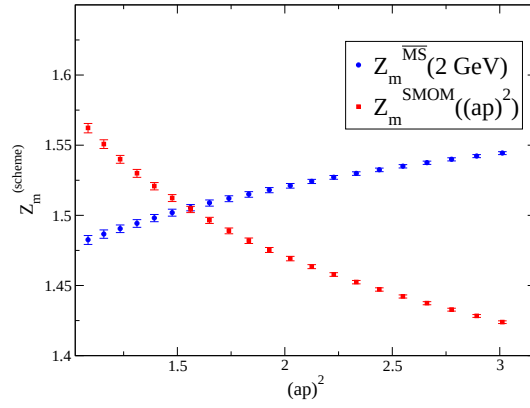
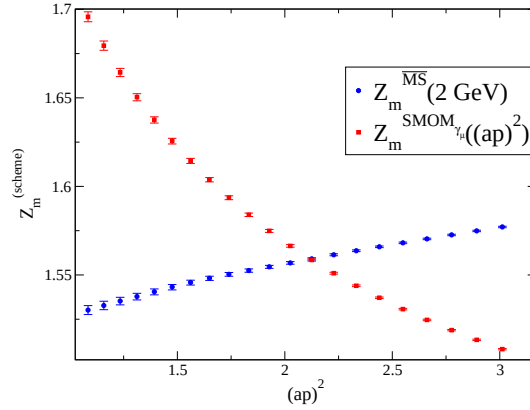
(a) $Z_m^{RI/MOM}((ap)^2) \rightarrow Z_m^{\overline{MS}}(2 \text{ GeV})$ (b) $Z_m^{SMOM-\not{p}}((ap)^2) \rightarrow Z_m^{\overline{MS}}(2 \text{ GeV})$ (c) $Z_m^{SMOM-\gamma_\mu}((ap)^2) \rightarrow Z_m^{\overline{MS}}(2 \text{ GeV})$

Figure 4.10: Z_m on the $\beta = 2.13$ lattice in a MOM scheme (red) and with the perturbative running divided out and converted to $\overline{MS}$ (blue).

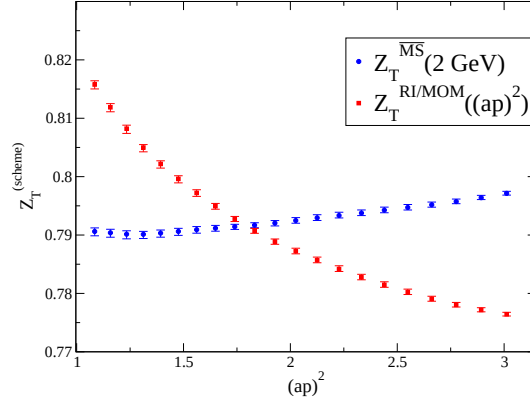
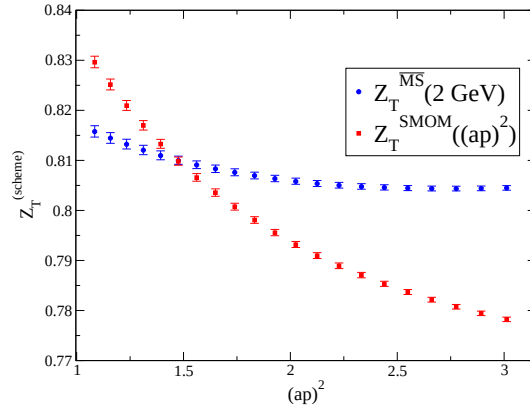
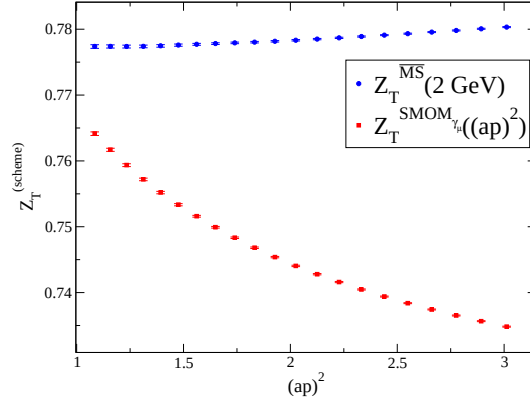
(a) $Z_T^{RI/MOM}((ap)^2) \rightarrow Z_T^{\overline{MS}}(2 \text{ GeV})$ (b) $Z_T^{SMOM-\not{p}}((ap)^2) \rightarrow Z_T^{\overline{MS}}(2 \text{ GeV})$ (c) $Z_T^{SMOM-\gamma_\mu}((ap)^2) \rightarrow Z_T^{\overline{MS}}(2 \text{ GeV})$

Figure 4.11: Z_T on the $\beta = 2.13$ lattice in a MOM scheme (red) and with the perturbative running divided out and converted to $\overline{MS}$ (blue).

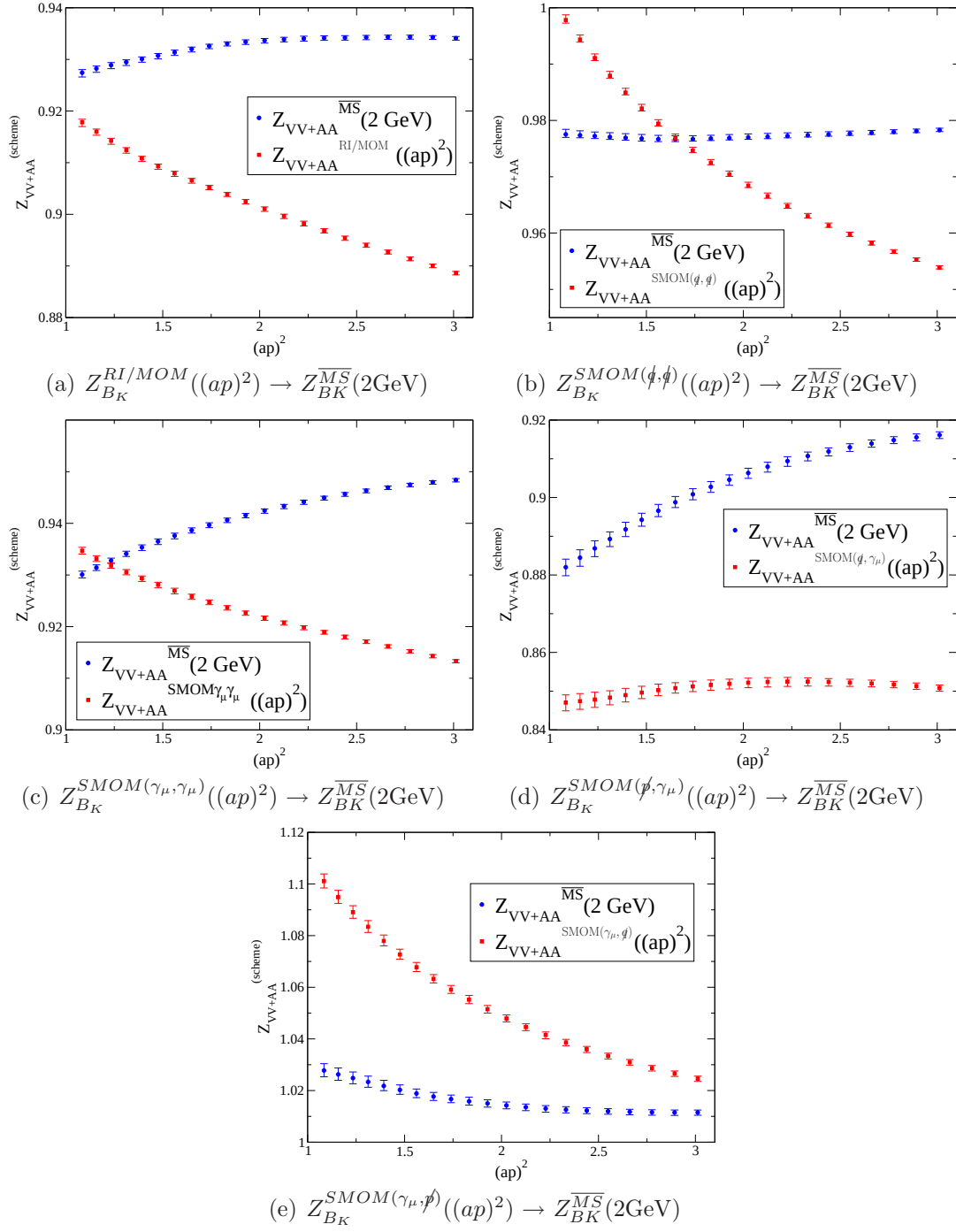


Figure 4.12: Z_{B_K} on the $\beta = 2.13$ lattice in a MOM scheme (red) and with the perturbative running divided out and converted to $\overline{MS}$ (blue).

at scale μ to $\overline{MS}$ at μ' we simply divide by $Z_O^{(S)}(\mu)$ and multiply by $Z_O^{(\overline{MS})}(\mu')$ which requires us to evaluate the integral above. To third order it is [38],

$$\begin{aligned}
c(x) &= \exp \left(\int^{\alpha(\mu)} \frac{\gamma(x)}{\beta(x)} dx \right) \\
&= (\alpha)^{\bar{\gamma}_0} [1 + (\bar{\gamma}_1 - \bar{\beta}_1 \bar{\gamma}_0) \alpha \\
&\quad + \frac{1}{2} [(\bar{\gamma}_1 - \bar{\beta}_1 \bar{\gamma}_0)^2 + \bar{\gamma}_2 + \bar{\beta}_1^2 \bar{\gamma}_0 - \bar{\beta}_1 \bar{\gamma}_1 - \bar{\beta}_2 \bar{\gamma}_0 \bar{\gamma}_0] \alpha^2 \\
&\quad + [\frac{1}{6} (\bar{\gamma}_1 - \bar{\beta}_1 \bar{\gamma}_0)^3 + \frac{1}{2} (\bar{\gamma}_1 - \bar{\beta}_1 \bar{\gamma}_0) (\bar{\gamma}_2 + \bar{\beta}_1^2 \bar{\gamma}_0 - \bar{\beta}_1 \bar{\gamma}_1 - \bar{\beta}_2 \bar{\gamma}_0 \\
&\quad + \frac{1}{3} (\bar{\gamma}_3 - \bar{\beta}_1^3 \bar{\gamma}_0 + 2 \bar{\beta}_1 \bar{\beta}_2 \bar{\gamma}_0 \\
&\quad - \bar{\beta}_3 \bar{\gamma}_0 + \bar{\beta}_1^2 \bar{\gamma}_1 - \bar{\beta}_2 \bar{\gamma}_1 - \bar{\beta}_1 \bar{\gamma}_2)] \alpha^3 \\
&\quad + O(\alpha^4)].
\end{aligned} \tag{4.123}$$

We use the highest order perturbation theory available for each of the schemes. For the RI-MOM scheme bilinears we use the three and four loop results of [38] and [77], for the RI-SMOM and RI-SMOM $_{\gamma_\mu}$ bilinears we use the three loop results in [76, 3, 78]. For BK we use the two loop perturbative results in [10]. We also need the beta function to evolve α to the scale of interest, that is we know alpha at a high scale and solve $\frac{d \log(\alpha)}{d \mu} = \beta$ to get α at a lower scale. For this we use the $\overline{MS}$ beta function of [38]. Although in general only the first two coefficients are scheme independent, the four loop RI β function is equivalent to the four loop $\overline{MS}$ β function, [77]. In this work we have only used perturbative results that have been written in terms of $\alpha^{\overline{MS}}$. Any change in the higher co-efficients from $\alpha^{RI} \rightarrow \alpha^{\overline{MS}}$ has already been reabsorbed. We start our coupling constant at $\alpha(M_Z) = 0.1176$ [4] use the five flavour beta function down to M_b , the four flavour down to M_c and the three flavour to run to the matching scale we want to renormalize at, usually $2GeV$ or $3GeV$. This procedure gives us renormalization constants in the $\overline{MS}$ scheme at the matching scale. On these smaller lattices we do not go as high as $3 GeV$ and renormalize instead at $2GeV$, which is the blue data in figures 4.9, 4.10, 4.11 and 4.12.

Of course this assumes that the running is exactly described by low order perturbation theory even at small momentum, where the coupling is large. This is obviously not the case as can be seen from the fact that the $\overline{MS}$ renormalization constants, with the running divided out, are not just constants but have some

curvature as a function of $(ap)^2$. Previously this had been all attributed to lattice artefacts and a linear extrapolation in $(ap)^2$ was taken by fitting the data in the high momentum region to a line and extrapolating to zero. This is a flawed procedure and mixes perturbative truncation errors with errors due to lattice artefacts. We will later take a continuum limit at non-zero momentum, eliminating discretization errors and isolating the perturbative error. However, as a check, we follow this flawed procedure after reference [8] and perform a linear fit in $(ap)^2$ in the range $2 < (ap)^2 < 3.2$ and obtain the results of tables 4.4 and 4.5 which compare well to the results of [8]. These numbers demonstrate already that momentum sources with twisted boundary conditions enable very high precision calculations of renormalization constants to be made very cheaply. Our approach instead of extrapolating is to take the point where $(ap) = \mu = 2\text{GeV}$ and convert from RI to $\overline{MS}$. We report these values also in tables 4.4 and 4.5.

Z_O	Ref: [8]	RIMOM (extrap)	
Z_q	0.7726(48 + 150)	0.7753(15)	
Z_m	1.656(30 + 83)	1.483(28)	
Z_T	0.7950(34 + 150)	0.7962(13)	
Z_O	RIMOM	SMOM- $\not{q}$	SMOM- γ_μ
Z_q	0.79605(63)	0.8035(10)	0.77744(31)
Z_m	1.546(15)	1.5073(26)	1.5405(18)
Z_T	0.80339(62)	0.80626(73)	0.77768(20)

Table 4.4: The quark field, mass and tensor current renormalization constants in $\overline{MS}$ at $2(\text{GeV})$. Error in the first column, top table is (stat + sys) error in the other columns statistical only. Note this work used 20 configurations at each mass whereas [8] used 300 point source measurements on 75 configurations. The second column, top table, uses the value extrapolated to zero for comparison with the results of [8]. The bottom table use simple interpolation to obtain the value at $p^2 = 2\text{GeV}$.

Ref: [8]	RIMOM (extrap)	RIMOM	
0.9276(52 + 220)	0.93330(73)	0.92994(54)	
SMOM- (g, g)	SMOM- (γ_μ, γ_μ)	SMOM- (g, γ_μ)	SMOM- (γ_μ, g)
0.97737(59)	0.93406(60)	1.0233(21)	0.8903(19)

Table 4.5: BK renormalization constant with same parameters as above. The extrapolated RIMOM is for comparison with [8] the rest of the measurements use the interpolated value at $p^2 = 2GeV$.

4.6.2 Scale determination in a small volume

Determining the lattice spacing for our $\beta = 2.23$ ensemble is a useful test case for the methods of section 4.5.2. To do so we have calculated r_C on our 16^3 lattices. We compute timelike Wilson loops with four hits of APE smearing, smearing parameter 2.5, in the spatial direction. The tree level improvement of [106] is required for the Iwasaki gluon action. We compute the tree level Wilson loop using code developed in [14] to obtain Table 4.6 where the potential includes the self energy part.

$$aV(\vec{R}a) = - \lim_{T \rightarrow \infty} \frac{1}{W(\vec{R}, T)} \frac{dW(\vec{R}, T)}{dT} = C_F V(\vec{R}) g^2 + O(g^4) \quad (4.124)$$

$\vec{R}$	V_W	V_I
(1,0,0)	0.166667	0.08963(1)
(2,0,0)	0.209842	0.12569(2)
(3,0,0)	0.225186	0.13879(3)
(4,0,0)	0.232442	0.14506(4)
(5,0,0)	0.236630	0.14878(4)
(6,0,0)	0.239366	0.15133(5)
(7,0,0)	0.241300	0.15318(5)
(8,0,0)	0.242742	0.15456(5)

Table 4.6: V_W is the static potential tree level part with the Wilson gluon action, V_I uses the Iwasaki gluon action. As R tends to infinity the difference between successive terms for Wilson and Iwasaki actions (the force) is the same since the self energy part cancels and the two actions reproduce the same IR physics.

By demanding

$$F(r_I) = (V(r) - V(r - a))/a = F_{tree} = \frac{1}{4\pi r_I^2} \quad (4.125)$$

we solve for r_I using Table 4.6. This approach leads to much reduced $O(a^2)$ errors [106] [80]. Further, using the force instead of the potential removes the self energy part and reduces the linear term, σr , to a constant. Fitting the Cornell potential to find r_C requires σ . However, σ is dependant on the long distance behaviour of $V(r)$ and, to constrain it, one has to sample large r . When fitting the force r_C can be obtained without including large distance data in the procedure and r_C computed this way is a finite volume safe observable with which to determine the lattice spacing.

In order to extract $V(r)$ from the average Wilson loop $\langle W(r, t) \rangle$ we first plot the effective potential,

$$\log\left(\frac{\langle W(r, t+1) \rangle}{\langle W(r, t) \rangle}\right), \quad (4.126)$$

as a function of t . The largest value of r used in this analysis is $4a$. Excited state contamination and statistical noise are problematic, and we use the black box method, see [61] and section 3.4.1, to define an improved effective mass that takes account of the first excited state. This allows us to extend t_{min} to $t = 3$ which gives a significant error reduction compared to the usual effective mass. On the $\beta = 2.13$ lattice we use 600 configurations for each mass, each configuration rotated to use 4 different time directions. For $\beta = 2.23$ we use 750 configurations per mass and again rotate the time direction.

The equation $r^2 F(r) = C$ has solutions in the range of our data for $C \in (0.6, 2.0)$. Lower values give more accurate r_C but these are likely to have large discretisation errors. For any C however $\frac{r_C/a_{2.13}}{r_C/a_{2.23}}$ should be constant, we plot this ratio as a function of C in figure 4.14 and determine $C > 1.4$ is appropriate and gives the value for the ratio of lattice scales,

$$R_a^2 = \frac{(a_{2.23})^2}{(a_{2.13})^2} = 0.652(21). \quad (4.127)$$

This implies a lattice spacing of $a^{-1} \simeq 2.14$ GeV for the $\beta = 2.23$ ensembles. We propose below a scale factor $s = 1.5$ and believe that at least two iterations of step scaling should be possible, based only on the static potential, without undue

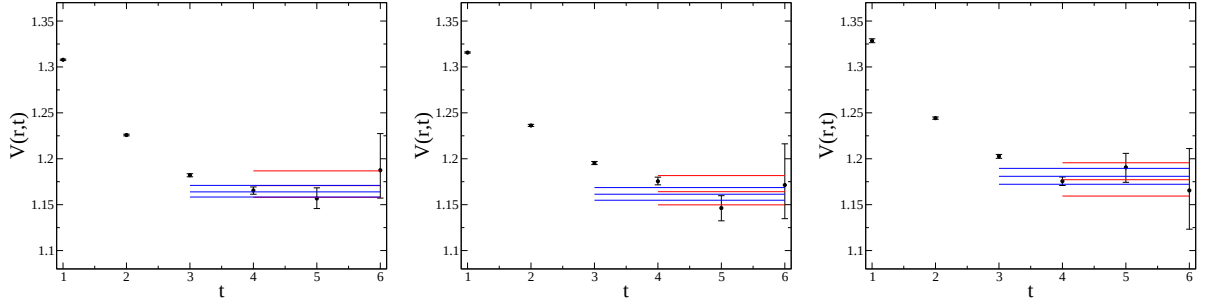


Figure 4.13: Effective mass plots for (left to right) $m = 0.03$, $m = 0.02$, $m = 0.01$ on the $\beta = 2.13$ lattice at $r = 4a$. Datapoints are computed via equation 4.126. The red constant is obtained from a fit over the last three datapoints. The blue is from the black box method with $t_{min} = 3$.

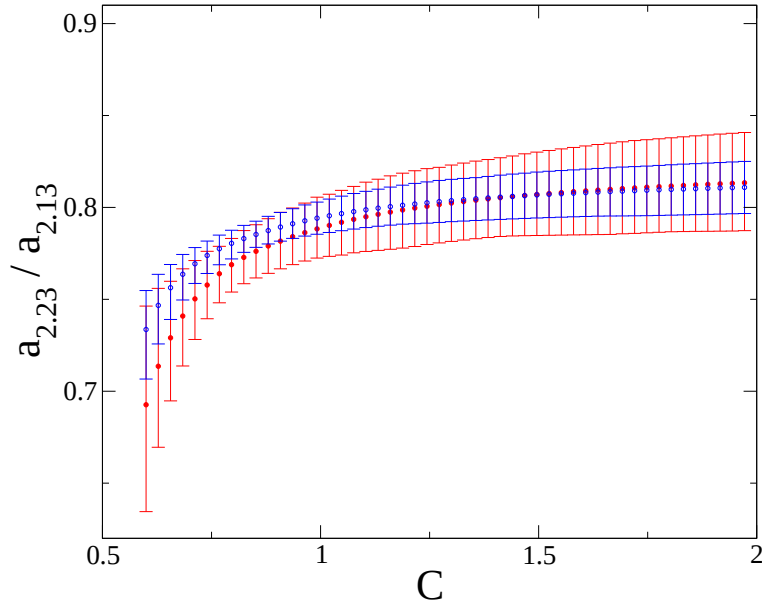


Figure 4.14: The ratio $\frac{r_C/a_{2.13}}{r_C/a_{2.23}}$ as a function of C . The red data is computed using the effective mass, equation 4.126, with $t_{min} = 4$ and the blue uses the black box method with $t_{min} = 3$, and gives the same value but with significantly smaller error.

error inflation. However, the static potential displays percent scale errors, even with many configurations, and finding a more precise alternative would certainly be good.

4.6.3 Scale Evolution Functions

We have two lattices one with $\beta = 2.13$ and lattice spacing $a^{-1} = 1.729\text{GeV}$ and one with $\beta = 2.23$ and lattice spacing $a^{-1} \simeq 2.14\text{GeV}$. The step scaling windows are given by, roughly, $0.35\text{GeV} \ll \mu \ll 5.5\text{GeV}$ on the $\beta = 2.13$ lattice and $0.45\text{GeV} \ll \mu \ll 6.75\text{GeV}$ on the $\beta = 2.23$ lattice. We can compute $\sigma_O(\mu, s\mu)$ using the two lattice spacings. We consider the evolution of renormalisation constants Z_q , Z_m , Z_T , and Z_{B_K} with both exceptional and the non-exceptional kinematics. We choose $\mu = 3(\text{GeV})$ and compute $\sigma_O(s\mu, \mu)$ as a function of s . Because of the very smooth μ dependence of the vertex functions computed with twisted boundary conditions we can perform a simple interpolation to match values of μ on each lattice. One could, in principle, simulate at the same physical momentum on each lattice by choosing the twisting angles appropriately, however, since the interpolation introduced only a very small uncertainty this was not necessary. For each operator we can now evaluate $\Sigma_O(s\mu, \mu = 3\text{GeV}, a^{-1} = 1.729\text{GeV})$ and $\Sigma_O(s\mu, \mu = 3\text{GeV}, a^{-1} = 2.14\text{GeV})$.

Linear extrapolation in a^2 with only two datapoints is not as robust as one might like, but we have found no obvious inconsistencies. We are helped in this regard by the off-shell automatic $O(a)$ improvement of domain wall fermions giving us smaller lattice artefacts. We do not observe large scaling violation when comparing the $\beta = 2.13$ to the $\beta = 2.23$ lattices for most quantities other than Z_q , which is a special case. However even in this case we always find quite close agreement with perturbation theory after the continuum limit with most schemes and quantities differing by at most a few percent. An extra lattice spacing would remove any doubts however we do not have access to one and we use the available data to make continuum extrapolations that seem very consistent with expectations.

Using these two lattice spacings, we can obtain the continuum limit step scaling function $\sigma_O(s\mu, \mu = 3\text{GeV})$ for the quark field (figure 4.15), mass (figure 4.16), tensor current (figure 4.17), and the four quark operator $\mathcal{O}_{VV+AA}$ (figure 4.18). With an appropriate third and smaller lattice spacing, using a

correspondingly smaller volume, we could similarly determine the next step, evolving from 3 GeV to 4.5 GeV and so on.

Figure 4.15 is especially important: Z_q suffers from the greatest lattice artefacts for domain wall fermions and is carefully eliminated in the NPR analysis of other operators. In the γ_μ scheme, with both exceptional and non-exceptional momenta, at finite lattice spacing the running of Z_q is in the *opposite* direction to the perturbative prediction. However the running recovered in the continuum limit is close to the perturbative prediction. The determination of Z_q with domain wall fermions displays the momentum dependence of the exponent, $\alpha(p)$, of localisation in the fifth dimension [120, 5]. As shown here, in a number of schemes, as long as an unambiguous continuum limit can be defined, the domain wall action will produce the continuum scaling behaviour of Z_q and the other renormalization constants. We also note the danger of doing the perturbative conversion without taking the continuum limit. In some cases, eg. Z_{BK} in the exceptional scheme, the agreement with lattice and perturbative results is better before the continuum limit is taken and might lead us to underestimate the error. We can also use this to identify good and bad schemes, for example in $SMOM(p, \gamma_\mu)$ the agreement with perturbation theory is very much worse than in the $SMOM(p, p)$ scheme, so in the final conversion it is far better to go from $SMOM(p, p)$ to $\overline{MS}$ than it is to use any of the other schemes as intermediates.

4.7 Volume Dependence

To satisfy the step scaling window condition at the lower end requires choosing $\mu \gg \frac{\pi}{aL}$. For $\beta = 2.13$ $\frac{\pi}{aL}$ is roughly 0.4GeV . In this section we check the volume dependence of the correlation functions by comparing simulations on two $\beta = 2.13$, Iwasaki, $L_s = 16$, DWF lattices; one with volume $16^3 \times 32$ and the other with volume $24^3 \times 64$. These 24^3 lattices were reported on in [2] and extended in [10]. In figure 4.19 we show the bilinear vertex functions computed on 16^3 and 24^3 lattices at different masses in the chiral limit. First note that the difference is small even at the lower end of the momentum region. For the vector vertex the difference $\Lambda_{16^3} - \Lambda_{24^3} = 0.0065$ and $\frac{\Lambda_{16^3} - \Lambda_{24^3}}{\Lambda_{16^3} + \Lambda_{24^3}} = 0.003$ so the finite volume correction at the lowest momentum considered is 0.3%. This is the worst case scenario, at higher momenta the finite volume correction gets smaller

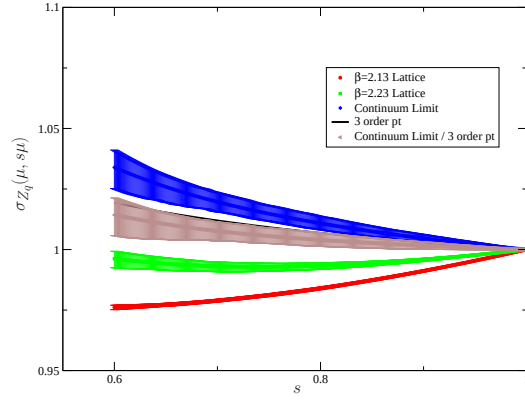
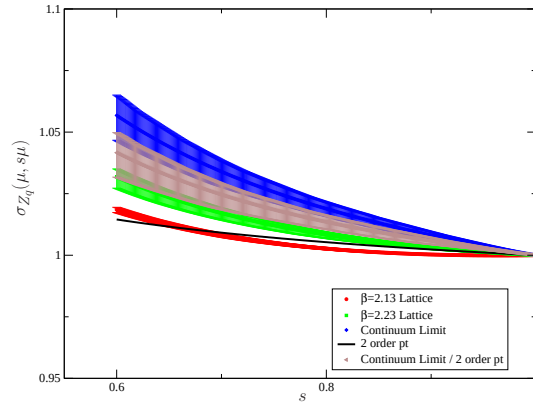
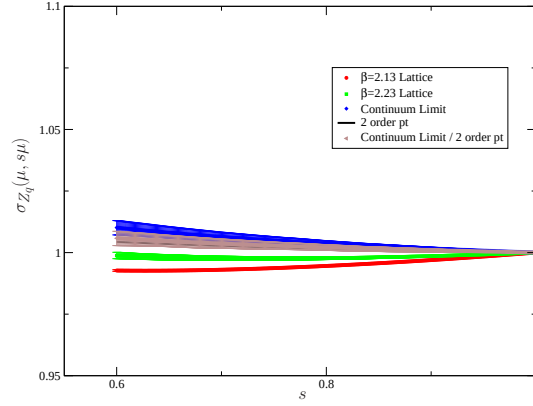
(a) Z_q , exceptional, RI (b) Z_q , non-exceptional, $SMOM_{\not{D}}$ (c) Z_q , non-exceptional, $SMOM_{\gamma_\mu}$

Figure 4.15: Quark field step scaling in the range $\simeq 2$ GeV to $\simeq 3$ GeV in the three different schemes compared to the scaling in each scheme.

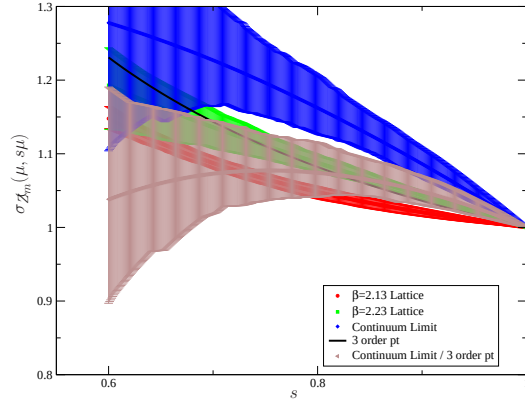
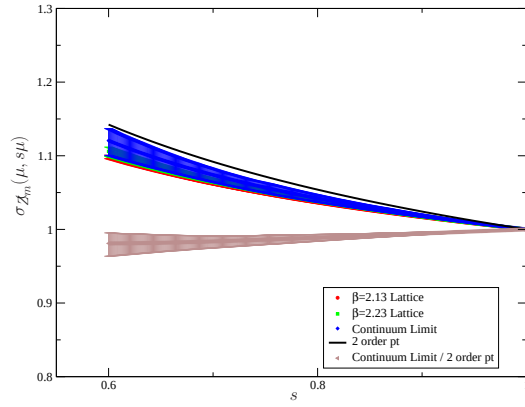
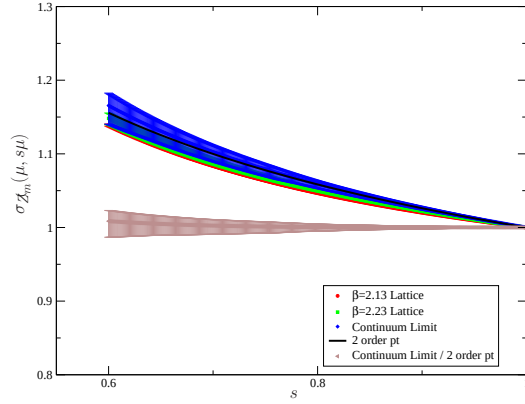
(a) Z_m , exceptional, RI (b) Z_m , non-exceptional, $SMOM_q$ (c) Z_m , non-exceptional, $SMOM_{\gamma_\mu}$

Figure 4.16: Quark mass step scaling in the range $\simeq 2$ GeV to $\simeq 3$ GeV in the three different schemes compared to the perturbative scaling in each scheme.

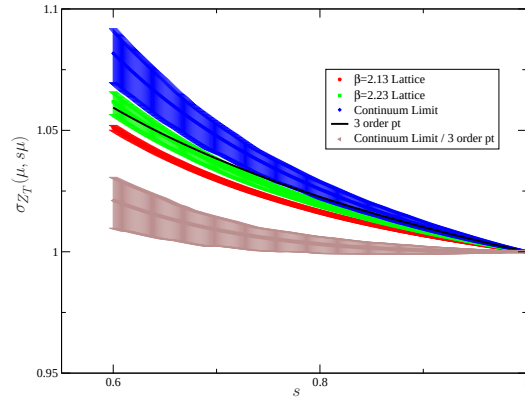
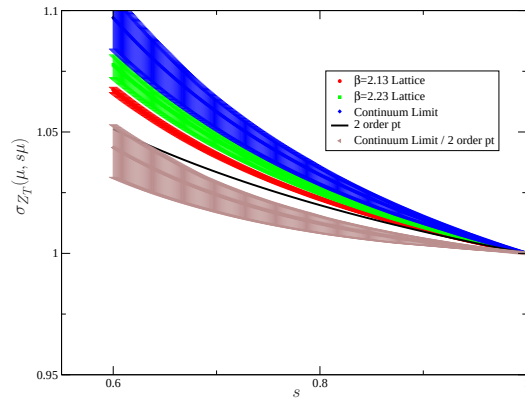
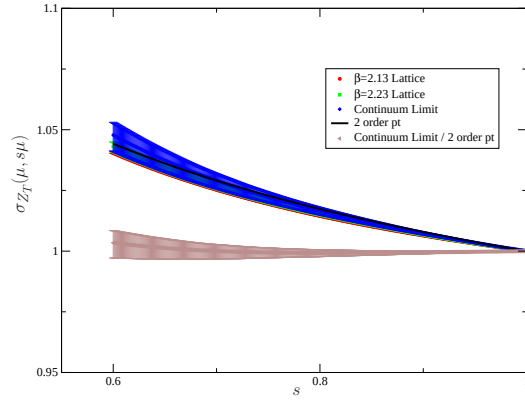
(a) Z_T , exceptional, RI (b) Z_T , non-exceptional, $SMOM_\phi$ (c) Z_T , non-exceptional, $SMOM_{\gamma_\mu}$

Figure 4.17: Tensor current step scaling in the range $\simeq 2$ GeV to $\simeq 3$ GeV in the three different schemes compared to the perturbative scaling in each scheme.

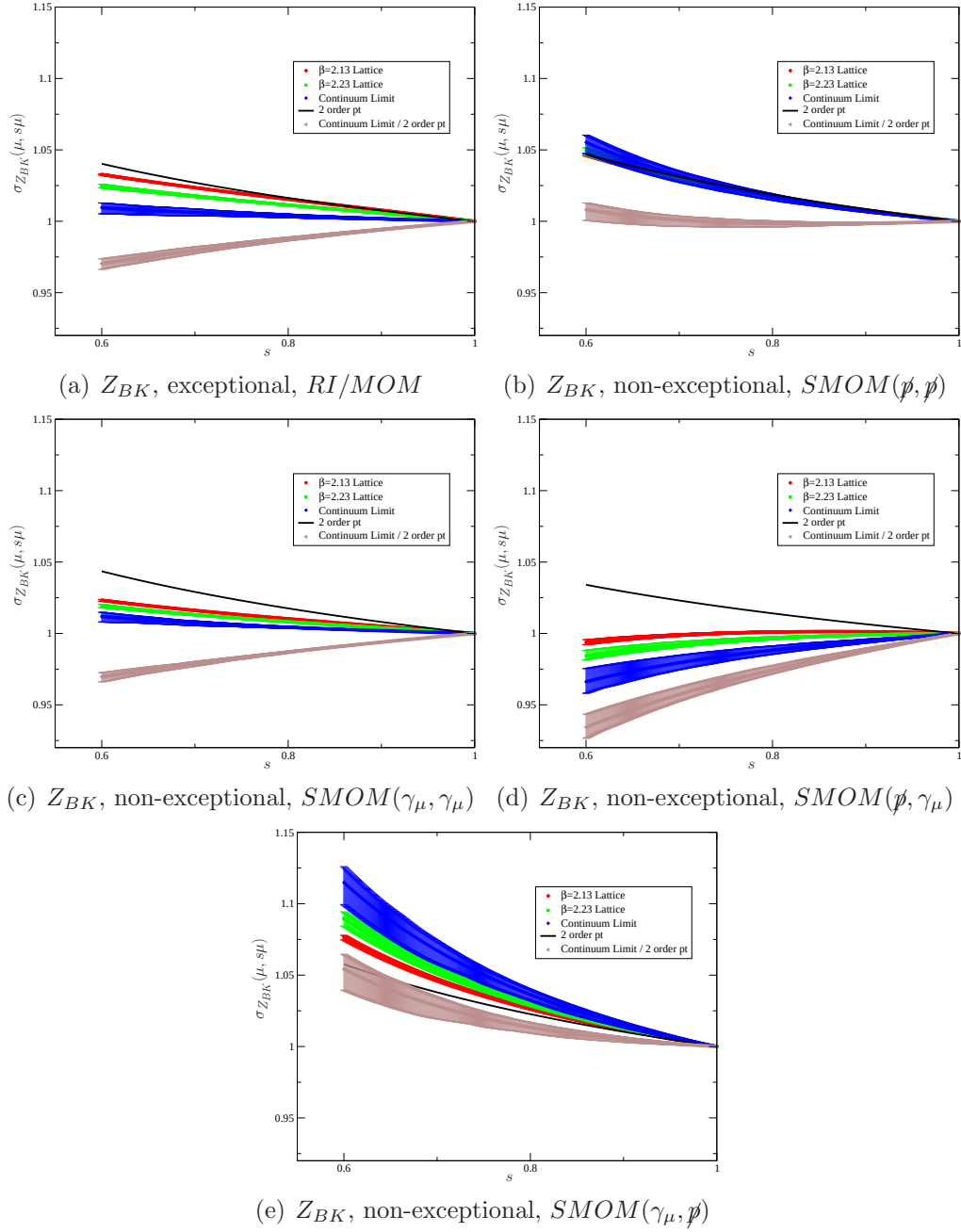


Figure 4.18: B_K renormalization running in the range $\simeq 2$ GeV to $\simeq 3$ GeV in the three different schemes compared to the perturbative scaling in each scheme.

and the other vertex functions have even smaller finite volume artefacts than the vector and axial. This is important because although we do not intend to use this method to scale to very large energies (say $O(100\text{GeV})$), we hope that three flavour step scaling up to around 5GeV is possible. The fact that the observed volume dependence is very weak means that we can operate in the small volumes necessary for this, as long as we satisfy the step scaling window condition, which at the lower end is quite a weak constraint.

4.8 Results on Large Lattices

β	Z_V	m_{res}	a^{-1}
2.25	0.7396(17)	0.0006664(76)	2.281(28)
2.13	0.7019(26)	0.003152(43)	1.729(25)

Table 4.7: Parameters from reference [10].

We also have access to larger volume lattices described in detail in reference [10] and used to renormalize B_K in reference [9]. These are $24^3 \times 64$ and $32^3 \times 64$, $L_s = 16$ dynamical domain wall lattices, generated with the Iwasaki gauge action, we refer to these as the 24^3 and 32^3 lattices in what follows. The quark masses on the 24^3 are $m_l = \{0.005, 0.01, 0.02, 0.03\}$ and $m_h = 0.04$ and on the 32^3 are $m_l = \{0.004, 0.006, 0.008\}$ and $m_h = 0.03$. Because the number of points in the volume is much larger, the benefits of volume sources are increased and the statistical error is reduced. Also the lattice parameters are much more accurately known, tabulated in 4.8, so we can obtain renormalization constants very accurately and calculate step scaling functions in the continuum limit with smaller errors than before. The momenta used are given in table 4.8 by specifying the base momentum and the twist angle, they span the physical range $1.1\text{GeV} < \mu < 3.3\text{GeV}$ on the 32^3 lattices and $1.1\text{GeV} < \mu < 3.0\text{GeV}$ on the 24^3 lattices. We point to our discussion from section 4.6.3 about extrapolations to the continuum with only two data points. Again, though a third β would be ideal we do not observe any inconsistencies and we now use a slightly finer lattice, $\beta = 2.25$, and again observe only a weak lattice spacing dependence, thanks to the $O(a)$ improvement of domain wall fermions.

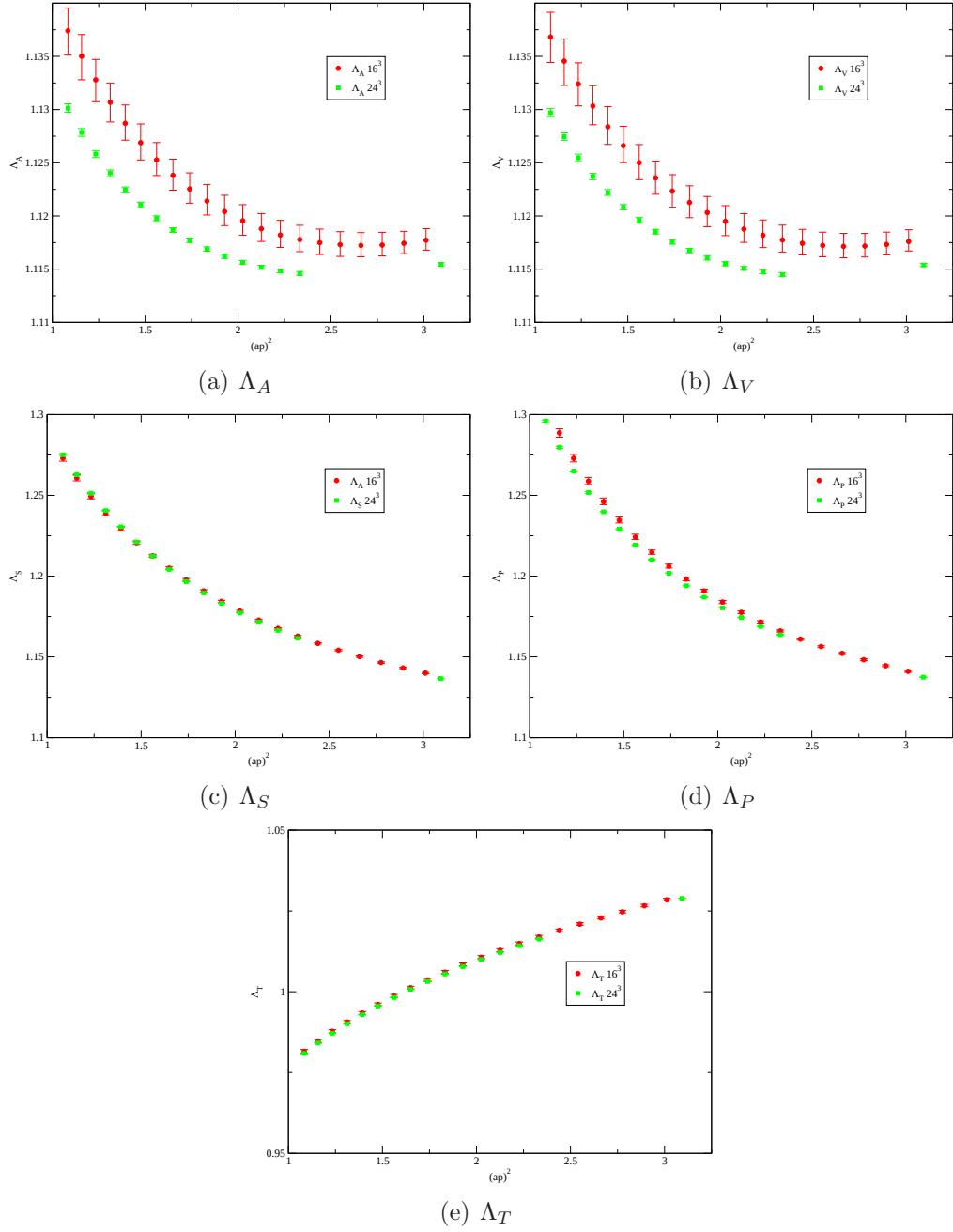


Figure 4.19: Λ_O for $O = \{S, P, V, A, T\}$ in the chiral limit comparing the 16^3 lattices (red) to the 24^3 lattices (green). Shown here is the SMOM scheme, the volume dependence in the other schemes is very similar.

$24^3 \times 64$	p_1	p_2	θ
	$(-2,0,2,0)$	$(0,2,2,0)$	$\{-0.45136, 0.732\}$
	$(-3,0,3,0)$	$(0,3,3,0)$	$\frac{3}{16}n : n = \{-2, 1, \dots, 12\}$
	$(-4,0,4,0)$	$(0,4,4,0)$	$\frac{3}{2}$
$32^3 \times 64$	p_1	p_2	θ
	$(-2,0,2,0)$	$(0,2,2,0)$	$\{-0.413, 0.783\}$
	$(-3,0,3,0)$	$(0,3,3,0)$	$\{\frac{1}{4}, 0.135\}$
	$(-4,0,4,0)$	$(0,4,4,0)$	$\{-\frac{3}{4}, \frac{3}{8}\}$
	$(-5,0,5,0)$	$(0,5,5,0)$	$\{-\frac{5}{8}, -0.531292, \frac{3}{8}\}$

Table 4.8: Non-exceptional momenta and twist angles used for the evaluation of amputated, twisted Green's functions in our NPR analysis. The momenta here are listed in (x, y, z, t) order for our 24^3 and 32^3 lattices. The integer Fourier mode numbers $\{n_i\}$ are related to the lattice momenta via $ap_i = \frac{n_i 2\pi}{L_i}$. The momentum added by the twist, B , is determined by the twist angle θ giving $ap_i = \frac{(2n_i + \theta)\pi}{L_i}$. The exceptional momenta used correspond to $p_2 = p_1$ for the same set of momenta. The twists that are not multiples of $\frac{1}{8}$ are chosen to match specific momenta on a larger volume lattice that will be described in a forthcoming publication.

4.8.1 Bilinear Renormalization

We tabulate the values of $Z_q^{\overline{MS}}(2, 3GeV)$, $Z_m^{\overline{MS}}(2, 3GeV)$ and $Z_T^{\overline{MS}}(2, 3GeV)$ obtained from the vertex functions on 24^3 and 32^3 lattices in tables 4.9, 4.10, 4.11. Momentum sources in large volumes give us great statistical precision. The sources of systematic error are:

a⁻¹: The lattice spacing is known up to some accuracy, on the 32^3 lattice $a^{-1} = 2.281(28)GeV$ on the 24^3 $a^{-1} = 1.729(25)GeV$. This affects the final result, since we measure renormalization constants at physical momentum, in GeV , converting from lattice units to physical ones introduces uncertainty. We have estimated the effect by performing the analysis twice, once with the central value and once with $a^{-1} + \Delta a^{-1}$ and quoted the difference as the systematic error.

VA: With exact chiral symmetry we have that $\Lambda_A = \Lambda_V$ however chiral symmetry is broken spontaneously and although highly suppressed by the non-exceptional kinematics there is still some slight difference. We perform the analysis twice, once using $\frac{1}{2}(\Lambda_A + \Lambda_V)$ to divide out Z_q and once using Λ_V and take the difference as a systematic. The results of tables 4.9, 4.10, 4.11 show that the effect is only important for exceptional momenta.

m_s: We perform the analysis in the limit where the light quark masses go to zero but we do not also extrapolate the heavy quark masses. To take account of this we use the slope found in extrapolating the light quarks and divide it by two to estimate $\frac{\partial Z}{\partial m}$. This times m_s is used to estimate the error due to the non-zero heavy quark mass. This is quite significant for the exceptional Z_m but small in most other cases.

SP: For Z_m using $\frac{1}{2}(\Lambda_S + \Lambda_P)$ or Λ_S is equivalent in the case of exact chiral symmetry. We use both and estimate the systematic as the difference. For the exceptional scheme the splitting between Λ_S and Λ_P is very large and the error is $O(50\%)$ from this source, even at large momentum.

Tables, 4.9, 4.10, 4.11 show the renormalization constants in $\overline{MS}$ calculated in each intermediate scheme. The systematic and statistical errors accounted for so far are very small, however, the results are not independent of the intermediate scheme due to perturbative truncation errors. We obtained the renormalization constants by fitting the data to the functional form,

$$f(x) = c_0 + \frac{c_{-1}}{x} + c_1 x \quad (4.128)$$

where $x = (ap)^2$. This is simply a good parametrization of the data over a small range. We use this to interpolate to the appropriate physical momentum. On 24^3 we use (the physical momentum required is on the left and the corresponding fit range used is on the right)

$$\begin{aligned} 2GeV : (ap)^2 &= [1.0, 2.5] \\ 3GeV : (ap)^2 &= [2.0, 3.2] \end{aligned} \quad (4.129)$$

on the 32^3 we use,

$$\begin{aligned} 2GeV : (ap)^2 &= [0.4, 1.4] \\ 3GeV : (ap)^2 &= [1.2, 2.5]. \end{aligned} \quad (4.130)$$

We test the interpolation by removing or adding one data point and find errors far below our already subdominant statistical error, so we neglect them.

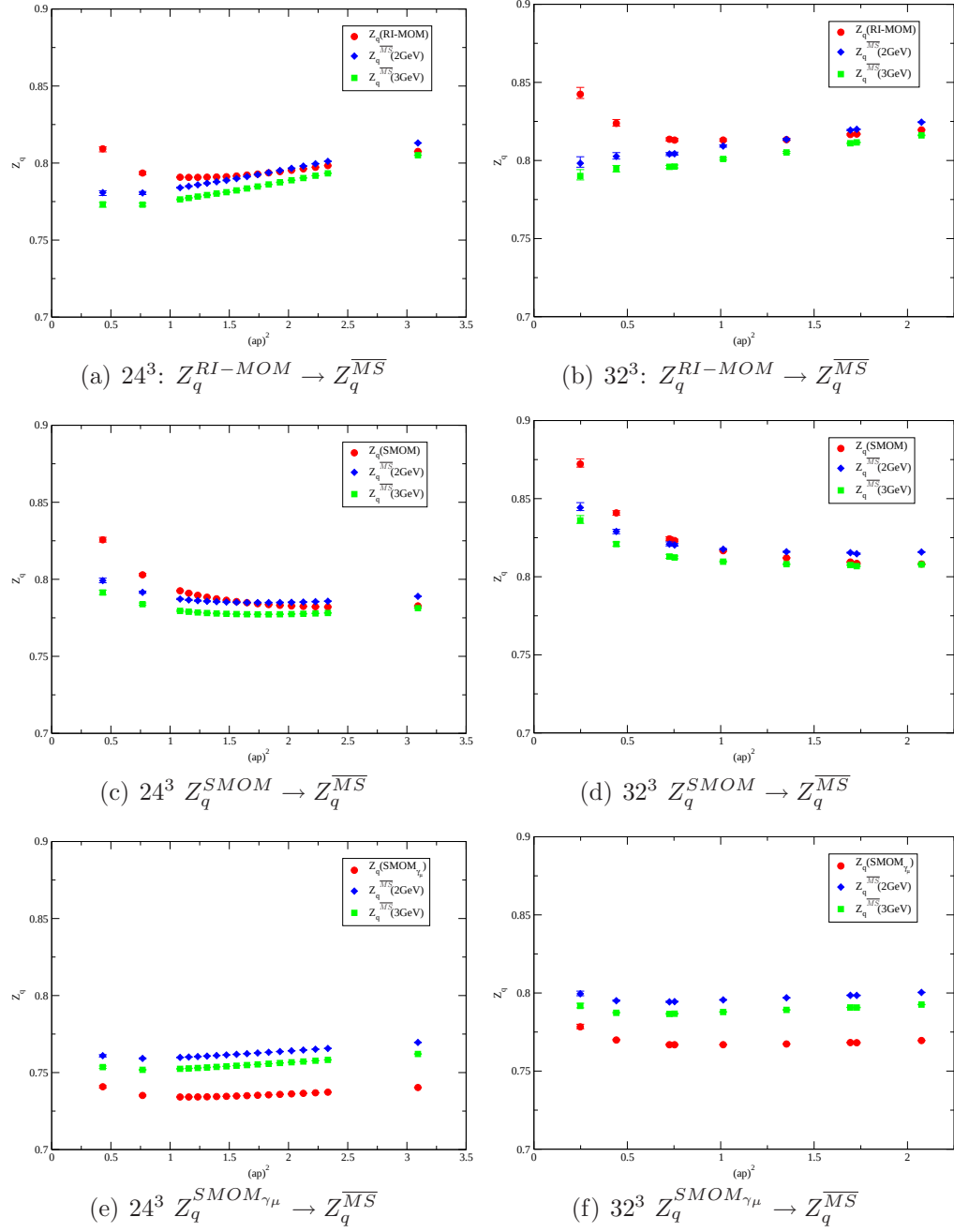


Figure 4.20: Z_q on the 24^3 (left) and 32^3 (right) lattices with the in RI-MOM (a), RI-SMOM (b), RI-SMOM $_{\gamma\mu}$ (c) in red. The perturbative running is then divided out and the conversion to $\overline{MS}$ is performed at 2GeV (blue) and 3GeV (green).

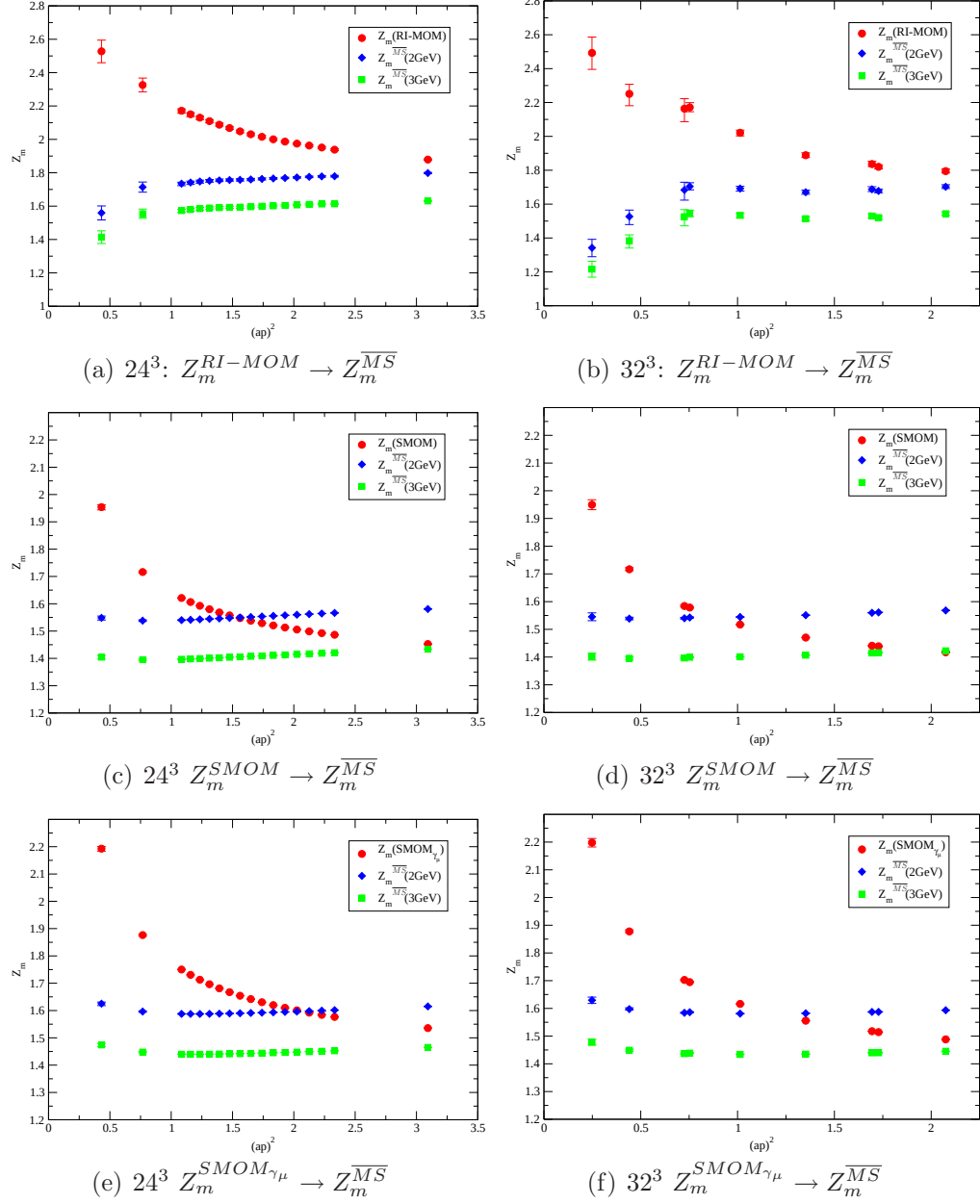


Figure 4.21: Z_m on the 24^3 (left) and 32^3 (right) lattices with the in RI-MOM (a), RI-SMOM (b), RI-SMOM $_{\gamma_\mu}$ (c) in red. The perturbative running is then divided out and the conversion to $\overline{MS}$ is performed at 2GeV (blue) and 3GeV (green).

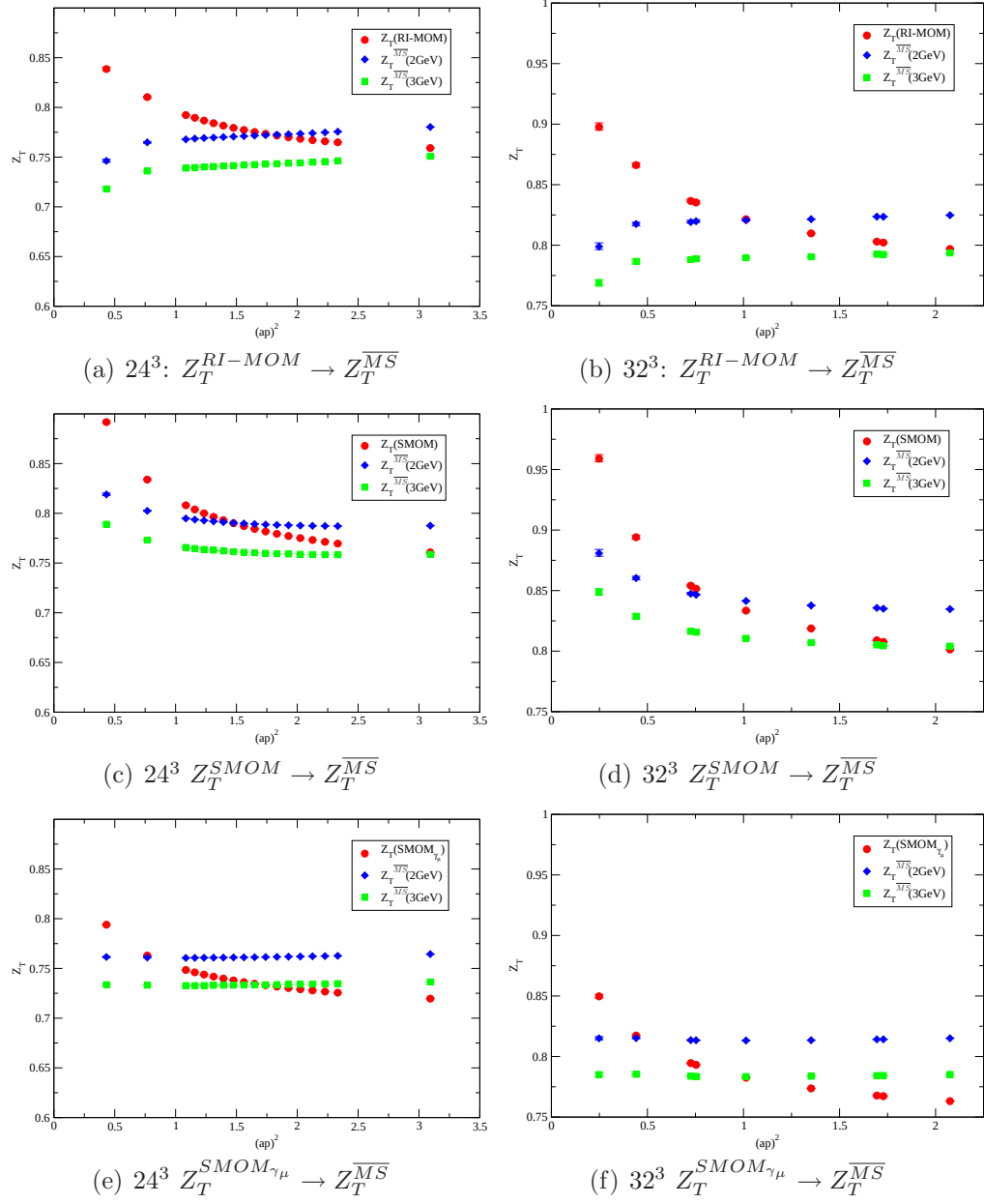


Figure 4.22: Z_T on the 24^3 (left) and 32^3 (right) lattices with the in RI-MOM (a), RI-SMOM (b), RI-SMOM $_{\gamma\mu}$ (c) in red. The perturbative running is then divided out and the conversion to $\overline{MS}$ is performed at 2GeV (blue) and 3GeV (green).

24 ³				32 ²		
	RI-MOM	RI-SMOM	RI-SMOM _{γ_μ}	RI-MOM	RI-SMOM	RI-SMOM _{γ_μ}
$Z_q^{\overline{MS}}(2.0GeV)$	0.78267	0.78560	0.76084	0.80507	0.82013	0.79457
stat	0.00024	0.00030	0.00013	0.00063	0.00056	0.00020
a^{-1}	0.00008	0.00054	0.00006	0.00021	0.00061	0.00006
m_s	0.00167	0.00359	0.00176	0.00274	0.00139	0.00074
VA	0.00300	0.00004	0.00005	0.00247	0.00005	0.00001
Total	0.00344	0.00364	0.00176	0.00375	0.00162	0.00077
Percent	0.44%	0.46 %	0.23 %	0.47%	0.20 %	0.10 %
	RI-MOM	RI-SMOM	RI-SMOM _{γ_μ}	RI-MOM	RI-SMOM	RI-SMOM _{γ_μ}
$Z_q^{\overline{MS}}(3.0GeV)$	0.80117	0.78085	0.76166	0.81158	0.80729	0.79083
stat	0.00013	0.00015	0.00009	0.00025	0.00028	0.00012
a^{-1}	0.00109	0.00017	0.00036	0.00035	0.00021	0.00014
m_s	0.00150	0.00205	0.00082	0.00036	0.00036	0.00003
VA	0.00089	0.00001	0.00003	0.00080	0.00001	0.00000
Total	0.00206	0.00206	0.00090	0.00098	0.00050	0.00018
Percent	0.26%	0.26 %	0.12 %	0.12%	0.06 %	0.02 %

Table 4.9: $Z_q^{\overline{MS}}(2GeV)$ and $Z_q^{\overline{MS}}(3GeV)$ on the 24³ and 32³ lattices using each intermediate scheme.

Step Scaling Functions

The lattice spacing is known more accurately on these lattices and we have smaller statistical errors and a larger scaling window, thus we expect that the step scaling functions can be obtained to much greater precision using the 24³ and 32³ lattices. We plot $\sigma_{\mathcal{O}}(s\mu, \mu)$ in figures 4.23, 4.24 and 4.25 and compare with perturbation theory, $\mu = 3GeV$ here and we have $s < 1$, so σ must equal one at 3GeV. The continuum limit step scaling function matches perturbation theory closely for some schemes and less so for others. Again we see that only in the continuum limit does comparison of non-perturbative and perturbative results make sense, again for Z_q the direction of the running changes after taking a continuum limit. For Z_m in the exceptional scheme there is surprisingly good agreement with perturbation theory on the 24³ lattice, which is shown to be spurious when extrapolated to the continuum. Schemes which match most closely their perturbative expansions are: Z_q - SMOM_{γ_μ}, Z_m - SMOM, Z_T - SMOM_{γ_μ}. We also provide tables, 4.14, of step scaling functions for each bilinear from 2GeV to 3GeV. The systematic errors are obtained as before.

24 ³				32 ²		
	RI-MOM	RI-SMOM	RI-SMOM _{γ_μ}	RI-MOM	RI-SMOM	RI-SMOM _{γ_μ}
$Z_m^{\overline{MS}}(2.0GeV)$	1.65593	1.54507	1.58807	1.69007	1.54131	1.58398
stat	0.00676	0.00062	0.00041	0.01369	0.00133	0.00100
a^{-1}	0.00784	0.00561	0.00701	0.00662	0.00520	0.00646
m_s	0.01778	0.00115	0.00363	0.11931	0.00201	0.00012
VA	0.00647	0.00009	0.00013	0.00571	0.00010	0.00005
SP	0.82233	0.00697	0.00715	1.09934	0.00488	0.00496
Total	0.82261	0.00904	0.01066	1.10592	0.00753	0.00821
Percent	49.68%	0.59 %	0.67 %	65.44%	0.49 %	0.52 %
	RI-MOM	RI-SMOM	RI-SMOM _{γ_μ}	RI-MOM	RI-SMOM	RI-SMOM _{γ_μ}
$Z_m^{\overline{MS}}(3.0GeV)$	1.57452	1.43239	1.46346	1.52499	1.41556	1.43993
stat	0.00282	0.00020	0.00014	0.00752	0.00044	0.00022
a^{-1}	0.00470	0.00318	0.00362	0.00428	0.00294	0.00359
m_s	0.00698	0.00123	0.00119	0.03101	0.00074	0.00067
VA	0.00174	0.00001	0.00005	0.00120	0.00001	0.00000
SP	0.31661	0.00067	0.00069	0.38269	0.00026	0.00023
Total	0.31674	0.00348	0.00388	0.38404	0.00307	0.00367
Percent	20.12%	0.24 %	0.26 %	25.18%	0.22 %	0.25 %

Table 4.10: $Z_m^{\overline{MS}}(2GeV)$ and $Z_m^{\overline{MS}}(3GeV)$ on the 24³ and 32³ lattices using each intermediate scheme.

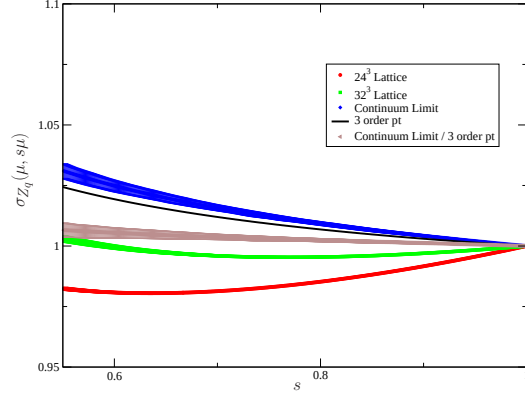
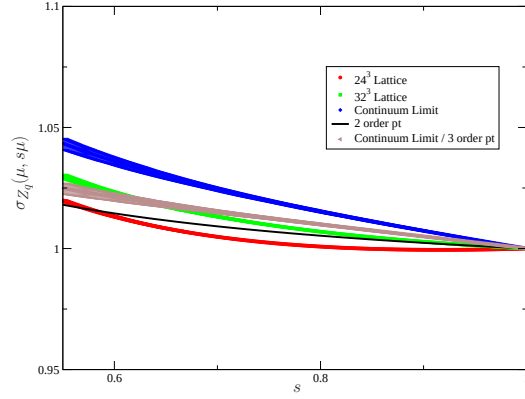
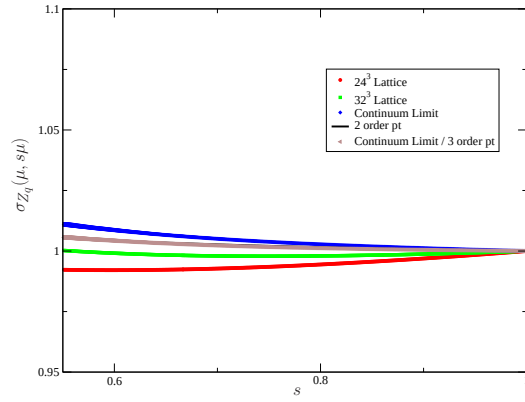
(a) Z_q , exceptional, RI (b) Z_q , non-exceptional, $SMOM$ (c) Z_q , non-exceptional, $SMOM_{\gamma_\mu}$

Figure 4.23: Quark field renormalization running from $s = 0.55$ to $s = 1$, where $\mu = 3\text{GeV}$ in three different schemes compared to the perturbative running in each scheme. Red, is Σ on the 24^3 lattice, green, Σ on the 32^3 lattice and blue the continuum limit σ . The solid black line is the highest order perturbation theory available and the brown data is $\frac{\sigma^{(non-pert)}}{\sigma^{(pert)}}$.

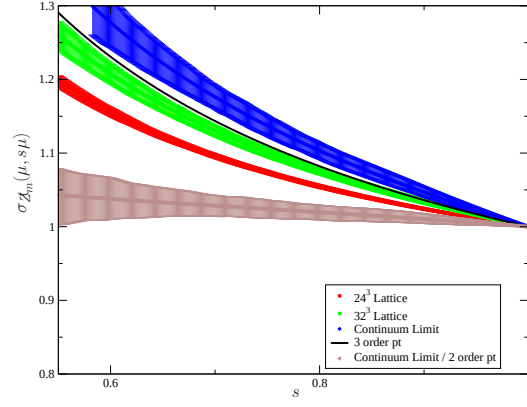
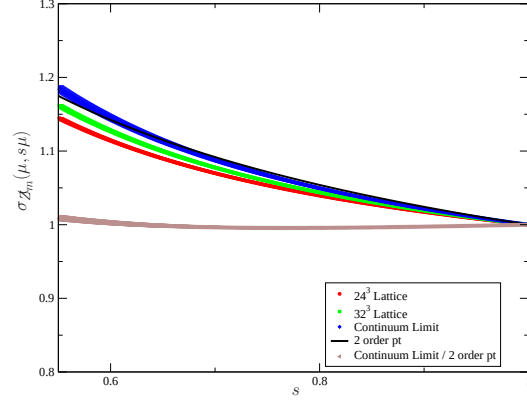
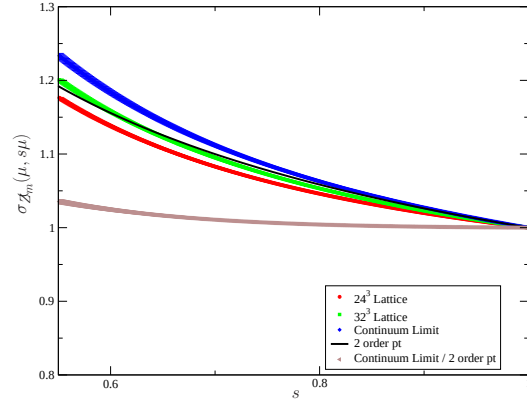
(a) Z_m , exceptional, RI (b) Z_m , non-exceptional, $SMOM$ (c) Z_m , non-exceptional, $SMOM_{\gamma_\mu}$

Figure 4.24: Quark mass renormalization running from $s = 0.55$ to $s = 1$, where $\mu = 3\text{GeV}$ in three different schemes compared to the perturbative running in each scheme. Red, is Σ on the 24^3 lattice, green, Σ on the 32^3 lattice and blue the continuum limit σ . The solid black line is the highest order perturbation theory available and the brown data is $\frac{\sigma^{(non-pert)}}{\sigma^{(pert)}}$.

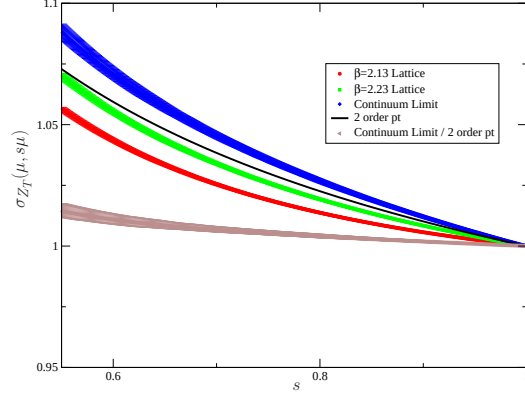
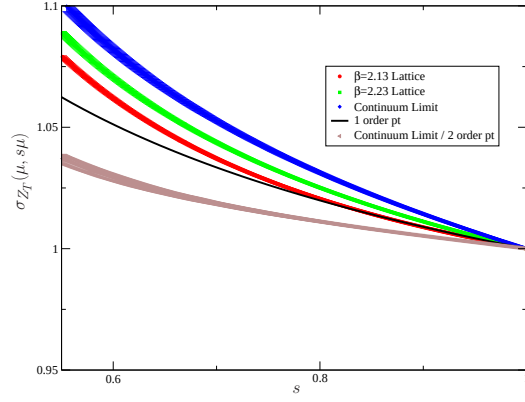
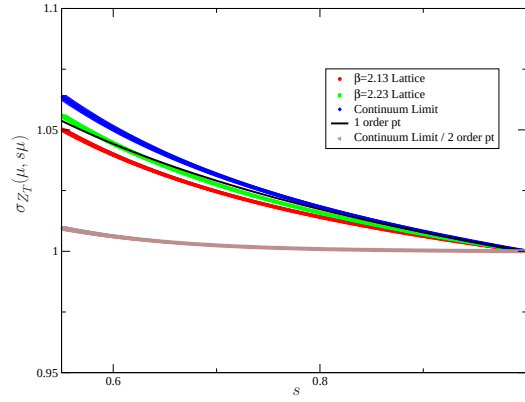
(a) Z_T , exceptional, RI (b) Z_T , non-exceptional, $SMOM$ (c) Z_T , non-exceptional, $SMOM_{\gamma_\mu}$

Figure 4.25: Tensor current renormalization running from $s = 0.55$ to $s = 1$, where $\mu = 3\text{GeV}$ in three different schemes compared to the perturbative running in each scheme. Red, is Σ on the 24^3 lattice, green, Σ on the 32^3 lattice and blue the continuum limit σ . The solid black line is the highest order perturbation theory available and the brown data is $\frac{\sigma^{(non-pert)}}{\sigma^{(pert)}}$.

24 ³				32 ²		
	RI-MOM	RI-SMOM	RI-SMOM _{γ_μ}	RI-MOM	RI-SMOM	RI-SMOM _{γ_μ}
$Z_T^{\overline{MS}}(2.0GeV)$	0.76972	0.79159	0.76084	0.81974	0.84615	0.81343
stat	0.00028	0.00023	0.00005	0.00060	0.00039	0.00011
a^{-1}	0.00118	0.00163	0.00098	0.00119	0.00160	0.00097
m_s	0.00362	0.00170	0.00061	0.00396	0.00049	0.00051
VA	0.00304	0.00003	0.00005	0.00268	0.00005	0.00004
Total	0.00488	0.00237	0.00116	0.00496	0.00172	0.00110
Percent	0.63%	0.30 %	0.15 %	0.61%	0.20 %	0.13 %
	RI-MOM	RI-SMOM	RI-SMOM _{γ_μ}	RI-MOM	RI-SMOM	RI-SMOM _{γ_μ}
$Z_T^{\overline{MS}}(3.0GeV)$	0.75033	0.75855	0.73619	0.79256	0.80477	0.78438
stat	0.00009	0.00009	0.00002	0.00026	0.00020	0.00005
a^{-1}	0.00051	0.00079	0.00058	0.00073	0.00096	0.00060
m_s	0.00081	0.00131	0.00004	0.00227	0.00043	0.00008
VA	0.00084	0.00001	0.00002	0.00069	0.00001	0.00000
Total	0.00127	0.00153	0.00058	0.00249	0.00107	0.00061
Percent	0.17%	0.20 %	0.08 %	0.31%	0.13 %	0.08 %

Table 4.11: $Z_T^{\overline{MS}}(2GeV)$ and $Z_T^{\overline{MS}}(3GeV)$ on the 24³ and 32³ lattices using each intermediate scheme.

Renormalized Quark Masses

We have calculated the quark mass renormalization constants in $\overline{MS}$ at higher scales than reference [10], where perturbation theory should be better behaved. Thus we hope to be able to improve on previous measurements of the quark mass. We perform an analysis of the light and strange quark masses following section VI of reference [10] using our twisted boundary condition renormalization data.

One estimate of the perturbative truncation error can be made by taking the conversion functions of [77, 38] and comparing the highest loop order to the lower ones.

$$C_m(S \rightarrow \overline{MS}, \mu) = 1 - \left(\frac{\alpha(\mu)}{4\pi} \right) C_1 - \left(\frac{\alpha(\mu)}{4\pi} \right)^2 C_2 - \dots = 1 - c_1 - c_2 - \dots \quad (4.131)$$

The two loop conversion represents a correction of $\frac{c_2}{1-c_1}\%$ and the three loop term,

c_3 , is unlikely to be larger. For each of the schemes we find this error estimate is,

$$\text{RI-MOM}(2\text{GeV}) : 5.9\% \quad (4.132)$$

$$\text{SMOM}_{\gamma_\mu}(2\text{GeV}) : 2.1\%$$

$$\text{SMOM}(2\text{GeV}) : 0.6\%$$

$$\text{RI-MOM}(3\text{GeV}) : 3.2\% \quad (4.133)$$

$$\text{SMOM}_{\gamma_\mu}(3\text{GeV}) : 1.4\%$$

$$\text{SMOM}(3\text{GeV}) : 0.4\%.$$

From this and also from the explicit comparison between perturbative and non-perturbative results in figure 4.24 we find that the SMOM scheme is the best for matching with perturbation theory. This is in contrast to reference [10], which used the SMOM_{γ_μ} to minimize the $\mathcal{H}(4)$ breaking scatter. This scatter is absent in our analysis and we are free to choose the scheme that agrees best with perturbation theory, so we use the SMOM results for our central values in the following analysis. Another estimate of the perturbative truncation error is to perform the analysis with two different intermediate schemes, SMOM and SMOM_{γ_μ} , and use the difference as the error [9], we will give this estimate below.

In order for several different ensembles to lie on the physical renormalized trajectory we can demand that ratios such as $\frac{m_K}{m_\Omega}$ and $\frac{m_\pi}{m_\Omega}$ are equal to their physical values. At different β s we will need different values of the bare light and strange masses to ensure these ratios equal their physical values. This defines functions $m_{ud}(\beta^e) = m_{ud}^e$ and $m_s(\beta^e) = m_s^e$. To compare bare quantities at different β s (ie. on different ensembles e) we can use ratios such as

$$Z_f^{(e)} = \frac{a^{(e)} m_f^{(M)}}{m_f^{(e)} a^{(M)}} \quad (4.134)$$

with $f = \{l, h\}$. These are the factors for the light and heavy quarks on ensemble e that relate them to the same scaling trajectory as a reference ensemble M . In our case there are only two ensembles, we choose the reference ensemble $M = 32^3$ thus $Z_f^{(M)} = 1$. The only non trivial factors are calculated in [10] table XXVI

and we take these over,

$$\begin{aligned} Z_l^{(24^3)} &= 0.981(9) \\ Z_h^{(24^3)} &= 0.974(7). \end{aligned} \tag{4.135}$$

Calculating these numbers involves chiral perturbation theory and the discussion of this would take us too far afield.

We extrapolate,

$$Z_{mf}^{\overline{MS}(e)c} = Z_m^{\overline{MS}(e)} / Z_f^{(e)} = Z_m^{\overline{MS}(e)} \frac{m_f^{(e)}}{a^{(e)}} \frac{a^{(M)}}{m_f^{(M)}} \tag{4.136}$$

as a function of the lattice spacing squared. The continuum limit of this is $\frac{m_f^{\overline{MS}}}{m_f^{(M)}}$. We finally multiply by $m_f^{(M)}$, which is obtained on the reference ensemble by using the physical omega baryon mass to set a physical scale, to obtain the required quantity $m_f^{\overline{MS}}$

Extrapolating $Z_{mf}^{\overline{MS}(e)}$ to the continuum using different schemes as intermediates we obtain,

$$\begin{aligned} Z_{ml}^{\text{SMOM}_{\gamma\mu} \rightarrow \overline{MS}(c)}(2GeV) &= 1.537(19) \\ Z_{mh}^{\text{SMOM}_{\gamma\mu} \rightarrow \overline{MS}(c)}(2GeV) &= 1.521(14) \\ Z_{ml}^{\text{SMOM} \rightarrow \overline{MS}(c)}(2GeV) &= 1.495(19) \\ Z_{mh}^{\text{SMOM} \rightarrow \overline{MS}(c)}(2GeV) &= 1.480(15) \end{aligned} \tag{4.137}$$

$$\begin{aligned} Z_{ml}^{\text{SMOM}_{\gamma\mu} \rightarrow \overline{MS}(c)}(3GeV) &= 1.370(17) \\ Z_{mh}^{\text{SMOM}_{\gamma\mu} \rightarrow \overline{MS}(c)}(3GeV) &= 1.355(13) \\ Z_{ml}^{\text{SMOM} \rightarrow \overline{MS}(c)}(3GeV) &= 1.355(17) \\ Z_{mh}^{\text{SMOM} \rightarrow \overline{MS}(c)}(3GeV) &= 1.341(13). \end{aligned} \tag{4.138}$$

The error in brackets is statistical only and comes mainly from the uncertainty in Z_l and Z_h which are included as fake bootstrap distributions with the appropriate mean and standard deviation. Using different intermediate schemes gives different continuum limit results. This is due to higher order perturbative effects and so the

difference between $Z_{ml}^{\overline{MS}(c)}$ calculated using different intermediate schemes gives an estimate of the perturbative truncation error;

$$\begin{aligned} Z_{mf}^{\overline{MS}(c)}(2GeV) &: 2.8\% \\ Z_{mf}^{\overline{MS}(c)}(3GeV) &: 1.1\%. \end{aligned} \tag{4.139}$$

The dependence on the intermediate scheme is very much reduced at $3GeV$. Renormalizing at the highest scale possible, once we have controlled lattice artefacts with a continuum limit, is the best approach and we renormalize at $3GeV$ instead of $2GeV$.

The plots in figure 4.24 demonstrate that SMOM agrees closely with perturbation theory and that this is the most appropriate for matching to $\overline{MS}$. This is supported by the small two loop term, c_2 , in the conversion function. We have found a non-perturbative scheme that is very close to $\overline{MS}$. Once we have found a scheme that agrees very closely with $\overline{MS}$ we should use it and we take the estimate of the perturbative truncation error from comparing high order terms in the conversion function within this scheme. To use the difference between $Z_{mf}^{\overline{MS}(c)}$ calculated via the two different intermediate schemes propagates the error from the less convergent scheme, $SMOM_{\gamma\mu}$, into the final result.

To obtain the final error we add in quadrature:

- the error in Z_V - 0.23%
- the systematic error from table 4.10 - 0.22%
- the perturbative error - 0.4%
- the statistical error in $Z_{mf}^{\overline{MS}(c)}$.

The last error contribution is the largest, it is $O(1\%)$. We then combine this error estimate with equation (64) in reference [10], the unrenormalized quark mass, to get:

$$\begin{aligned} m_{ud}^{\overline{MS}}(3GeV) &= 3.18(11)_{stat}(12)_{sys}(4)_{ren} \\ m_s^{\overline{MS}}(3GeV) &= 85.42(1.21)_{stat}(0.13)_{sys}(1.18)_{ren}. \end{aligned} \tag{4.140}$$

We can also quote a less accurate result at the more commonly used reference point of 2GeV for comparison with other lattice results,

$$\begin{aligned} m_{ud}(2\text{GeV}) &= 3.51(12)_{\text{stat}}(13)_{\text{sys}}(5)_{\text{ren}} \\ m_s(2\text{GeV}) &= 94.30(1.33)_{\text{stat}}(0.15)_{\text{sys}}(1.43)_{\text{ren}}. \end{aligned} \quad (4.141)$$

Compared to the analysis in [10] we have similar statistical and systematic errors, however, we have managed to half the renormalization error. This mostly comes from the choice of scheme and scale; the perturbative truncation error in SMOM is smaller than in SMOM_{γ_μ} and further reduced by using the higher scale. We consider the results above as improvement over the analysis in [10] because (a) we have accounted for the error in a^{-1} (b) we have taken a rigorous continuum limit without interpolating between $\mathcal{H}(4)$ inequivalent momenta and (c) we used an intermediate scheme that agrees better with perturbation theory as well as a higher scale. The mass at 2GeV can be obtained from the 3GeV number by using the reciprocal of the step scaling function in table 4.13 as a consistency check.

4.8.2 B_K Renormalization

In this section we describe the renormalization of B_K . This analysis was performed in reference [9] from which we take the four different schemes, equations 4.96, and the perturbation theory. The momenta used are the same in reference [9], they are given in table 4.8 as the twists that are multiples of $\frac{1}{8}$. From our experience with bilinears leaving out the extra momentum points will cause no loss in accuracy. The interpolations are very precise and introduce minimal error, we are dominated by systematic and perturbative errors. We use,

$$Z_{B_K} = \left(\frac{\Lambda_A^2}{\Lambda_{VV+AA}} \right) \quad (4.142)$$

removing factors of Z_q with the axial vector vertex function. In figure 4.26 we show the data on 24^3 and 32^3 for all five intermediate renormalization schemes converted to $\overline{MS}$ at 2GeV . The flattest data is observed for the “unmixed” schemes $\text{SMOM}(\not{q}, \not{q})$ and $\text{SMOM}(\gamma_\mu, \gamma_\mu)$, indicating that these schemes are closer to $\overline{MS}$.

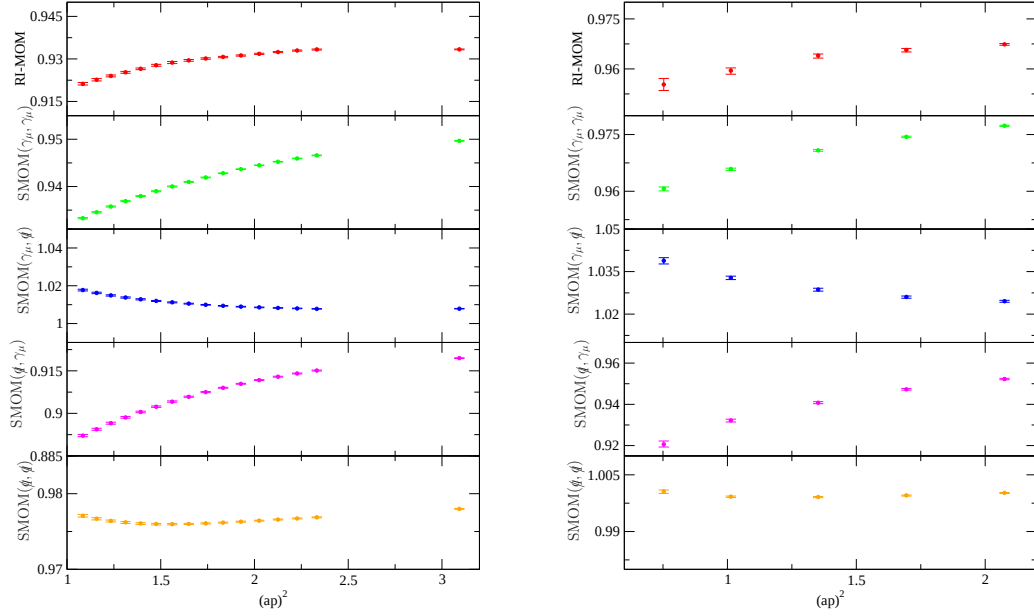


Figure 4.26: Using the perturbative running to convert the chiral limit of the ratio $Z_{B_K}^S$ to $Z_{B_K}^{MS}$ at 2 GeV for each p^2 . This is displayed for all five intermediate MOM schemes, 4.96, on the 24^3 (left) and 32^3 (right). The top two panels correspond to the original RI-MOM as the intermediate scheme and the other four rows correspond to the non-exceptional schemes.

We give results for $Z_{B_K}^{MS}$ at 2 and 3 GeV in each intermediate scheme in table 4.16 with several sources of uncertainty. Statistical, a^{-1} , chiral symmetry breaking (VA) and m_s error are calculated as in the previous section. We give the continuum limit step scaling from 2GeV to 3GeV in table 4.17 and plot the step scaling functions compared to perturbation theory in figure 4.27.

Because we do not have second order perturbative matching factors for B_K using the last term as an estimate of the perturbative truncation error is not as reliable. Instead we take the step scaling functions which are continuum limit, non-perturbative quantities and see how well they agree with perturbation theory in the momentum range we have simulated. In this case we find, as we found on smaller lattices, that the $\text{SMOM}(q, q)$ running is most closely described by its perturbative expansion. For an estimate of the error we take the difference between $\text{SMOM}(q, q)$ and $\text{SMOM}(\gamma_\mu, \gamma_\mu)$. We also note that, at the higher scale of 3GeV , which we can more confidently use because we intend to take a continuum

limit and remove discretization errors, the difference is much reduced. It is better then, as is the case for Z_m , to renormalize at 3GeV where we have a reduced systematic error from perturbation theory.

In the case of explicitly broken chiral symmetry $\mathcal{O}_{VV+AA}$ can mix with the operators $\mathcal{O}_{VV-AA}$, $\mathcal{O}_{SS\pm PP}$ and $\mathcal{O}_{TT}$ under renormalization. We show the magnitude of the mixing coefficients in figure 4.28 and see that they are significantly smaller than the multiplicative renormalization Z_{B_K} and can be safely neglected at our current level of precision.

Combining these renormalization constants with the analysis of reference [9] we obtain,

$$\begin{aligned} B_K^{\overline{MS}}(2\text{GeV}) &= 0.549(5)_{stat}(15)_\chi(2)_{FV}(21)_{ren} \\ B_K^{\overline{MS}}(3\text{GeV}) &= 0.529(5)_{stat}(15)_\chi(2)_{FV}(11)_{ren} \end{aligned} \quad (4.143)$$

The renormalization systematic is halved at 3GeV compared to 2GeV , a big improvement. This result can be converted to the renormalization group invariant B_K by dividing out $Z_{B_K}^{\overline{MS}}(3\text{GeV})$ to give,

$$B_K^{RGI} = 0.749(7)_{stat}(21)_\chi(3)_{FV}(15)_{ren} \quad (4.144)$$

which has a total error of 3.6%. Analyses from other groups since then support this result, see eg. [55], [88]. [12], [93].

4.9 Conclusion

We have introduced a new method, twisted boundary conditions to remove $\mathcal{H}(4)$ breaking artefacts in the calculation of vertex functions necessary for Rome-Southampton renormalization. Together with the statistical precision of volume sources and the chiral symmetry breaking suppression of non-exceptional kinematics we have an excellent method for calculating non-perturbative renormalization constants on the lattice. This combination of techniques also allows us to calculate step scaling functions from lattice QCD simulations in a statistically and theoretically clean way. Thus the Rome-Southampton window, long a source of difficulty in non-perturbative renormalization, has been opened. We do not try to remove perturbative and lattice artefacts on a single lattice, we deal with

them in turn. A continuum limit of the step scaling function or the renormalized quantity removes the $O(a^2)$ effects. We are then left with an object that runs with the momentum scale. All of this running can be attributed to the anomalous dimension of the operator. This is encoded in our step scaling functions, they quantify the running with momentum scale of the operator and are calculated fully non-perturbatively. Physical quantities can be calculated at low energy scales, where lattice artefacts are small and if we want them at higher scales we can use the step scaling functions. Perturbation theory only enters at the end when the quantity, renormalized at a high scale, is required in $\overline{MS}$. Because we are now able to reach high scales, perturbation theory will hopefully be better behaved and the conversion functions more accurate. As we have seen, going from 2GeV to 3GeV significantly reduced the error in Z_m and especially B_K .

We hope in future to calculate a database of step-scaling functions for a variety of QCD operators, since they are continuum limit quantities they are independent of the underlying action. We have also demonstrated that they may be obtained using very reasonable lattices, of $O(16^3)$, and heavy quark masses. Setting the scale seems to be the biggest challenge, however, new approaches to this, [28], are very promising. Step scaling and twisted boundary condition renormalization have become the standard in the RBC-UKQCD collaboration and future publications have and will continue to make much use of this new method.

	RI-MOM	RI-SMOM	RI-SMOM _{γ_μ}
$\sigma_{Z_q}(2GeV, 3GeV)$	0.98114	0.97254	0.99388
stat	0.00149	0.00089	0.00031
a^{-1}	0.00007	0.00019	0.00009
m_s	0.00189	0.00129	0.00039
VA	0.00220	0.00012	0.00049
Total	0.00326	0.00159	0.00070
Percent	0.33%	0.16 %	0.07 %

Table 4.12: Continuum limit step scaling function $\sigma_{Z_q}(2GeV, 3GeV)$.

	RI-MOM	RI-SMOM	RI-SMOM _{γ_μ}
$\sigma_{Z_m}(2GeV, 3GeV)$	0.82602	0.90543	0.88261
stat	0.01572	0.00157	0.00107
a^{-1}	0.00091	0.00050	0.00051
m_s	0.03311	0.00054	0.00343
VA	0.00387	0.00003	0.00009
SP	0.31819	0.00026	0.00014
Total	0.32032	0.00175	0.00363
Percent	38.78%	0.19 %	0.41 %

Table 4.13: Continuum limit step scaling function $\sigma_{Z_m}(2GeV, 3GeV)$.

	RI-MOM	RI-SMOM	RI-SMOM _{γ_μ}
$\sigma_{Z_T}(2GeV, 3GeV)$	0.94960	0.94210	0.96420
stat	0.00122	0.00077	0.00025
a^{-1}	0.00013	0.00001	0.00016
m_s	0.01076	0.00352	0.00310
VA	0.00167	0.00028	0.00035
Total	0.01096	0.00362	0.00313
Percent	1.15%	0.38 %	0.32 %

Table 4.14: Continuum limit step scaling function $\sigma_{Z_T}(2GeV, 3GeV)$.

24^3					
	RI-MOM	SMOM (γ_μ, γ_μ)	SMOM ($\gamma_\mu, \not{q}$)	SMOM ($\not{q}, \gamma_\mu$)	SMOM ($\not{q}, \not{q}$)
$Z_{B_K}^{\overline{MS}}(2.0GeV)$	0.92578	0.93731	1.01350	0.89936	0.97621
Stat	0.00028	0.00010	0.00032	0.00027	0.00011
a^{-1}	0.00049	0.00064	0.00225	0.00013	0.00140
m_s	0.00757	0.00393	0.00445	0.00054	0.00180
$V - A$	0.00750	0.00021	0.00026	0.00021	0.00026
Total	0.01067	0.00399	0.00500	0.00065	0.00230
Percent	1.15%	0.43 %	0.49 %	0.07 %	0.24 %
	RI-MOM	SMOM (γ_μ, γ_μ)	SMOM ($\gamma_\mu, \not{q}$)	SMOM ($\not{q}, \gamma_\mu$)	SMOM ($\not{q}, \not{q}$)
$Z_{B_K}^{\overline{MS}}(3.0GeV)$	0.90444	0.91983	0.97455	0.89147	0.94672
Stat	0.00066	0.00010	0.00029	0.00027	0.00011
a^{-1}	0.00076	0.00051	0.00131	0.00007	0.00084
m_s	0.00347	0.00181	0.00164	0.00148	0.00063
$V - A$	0.00203	0.00003	0.00012	0.00009	0.00012
Total	0.00415	0.00188	0.00213	0.00151	0.00106
Percent	0.46%	0.20 %	0.22 %	0.17 %	0.11 %

Table 4.15: $Z_{B_K}^{\overline{MS}}(2GeV)$ and $Z_{B_K}^{\overline{MS}}(3GeV)$ on the 24^3 lattice using each intermediate scheme.

32^3					
	RI-MOM	SMOM (γ_μ, γ_μ)	SMOM ($\gamma_\mu, \not{q}$)	SMOM ($\not{q}, \gamma_\mu$)	SMOM ($\not{q}, \not{q}$)
$Z_{B_K}^{MS}(2.0GeV)$	0.95541	0.96089	1.03838	0.92164	1.00028
Stat	0.00151	0.00046	0.00093	0.00104	0.00036
a^{-1}	0.00045	0.00052	0.00211	0.00030	0.00129
m_s	0.00846	0.00221	0.00386	0.00174	0.00151
$V - A$	0.00551	0.00014	0.00013	0.00010	0.00014
Total	0.01022	0.00232	0.00450	0.00205	0.00202
Percent	1.07%	0.24 %	0.43 %	0.22 %	0.20 %
	RI-MOM	SMOM (γ_μ, γ_μ)	SMOM ($\gamma_\mu, \not{q}$)	SMOM ($\not{q}, \gamma_\mu$)	SMOM ($\not{q}, \not{q}$)
$Z_{B_K}^{MS}(3.0GeV)$	0.93453	0.94284	0.99252	0.91681	0.96698
Stat	0.00030	0.00017	0.00034	0.00038	0.00013
a^{-1}	0.00058	0.00049	0.00137	0.00004	0.00086
m_s	0.00181	0.00048	0.00039	0.00024	0.00009
$V - A$	0.00188	0.00002	0.00002	0.00002	0.00002
Total	0.00269	0.00070	0.00147	0.00046	0.00088
Percent	0.29%	0.07 %	0.15 %	0.05 %	0.09 %

Table 4.16: $Z_{B_K}^{MS}(2GeV)$ and $Z_{B_K}^{MS}(3GeV)$ on the 32^3 lattice using each intermediate scheme.

	RI-MOM	SMOM (γ_μ, γ_μ)	SMOM ($\gamma_\mu, \not{q}$)	SMOM ($\not{q}, \gamma_\mu$)	SMOM ($\not{q}, \not{q}$)
$\sigma_{B_K}(2GeV, 3GeV)$	0.98457	0.98346	0.93783	1.00893	0.96189
stat	0.00352	0.00091	0.00154	0.00186	0.00073
m_s	0.01041	0.00075	0.00382	0.00056	0.00012
VA	0.00068	0.00066	0.00008	0.00042	0.00007
Total	0.01101	0.00135	0.00412	0.00199	0.00075
Percent	1.12%	0.14 %	0.44 %	0.20 %	0.08 %

Table 4.17: Scaling factor $\sigma_{B_K}(2\text{ GeV}, 3\text{ GeV})$ from 2 to 3 GeV for each scheme. The values are the reciprocal of the left most point in Figure 4.27. The error from the uncertainty in the lattice spacing is now folded into the statistical error.

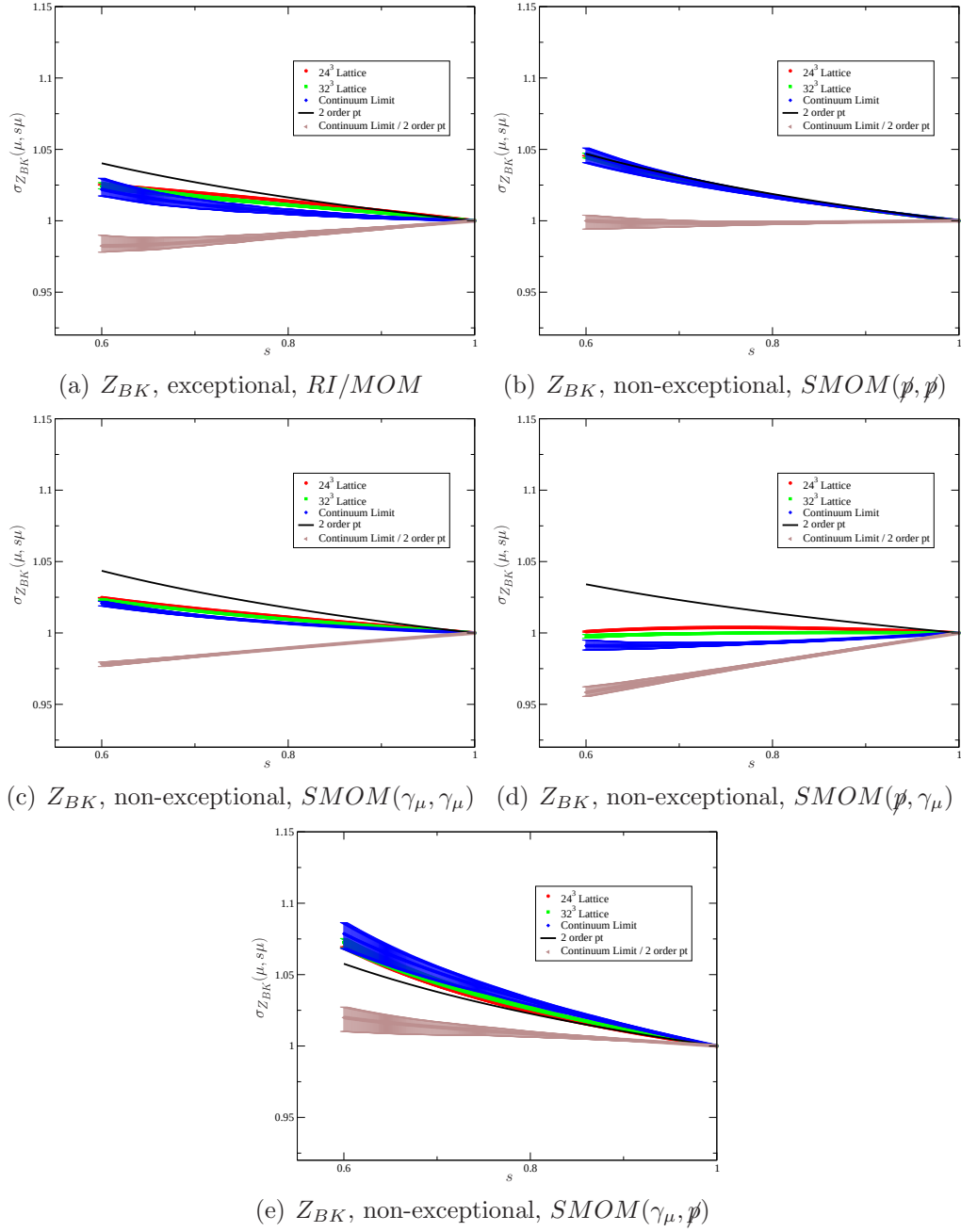


Figure 4.27: Continuum limit step scaling functions for all four SMOM and the RI-MOM scheme schemes (blue) compared with one-loop perturbation theory (black). The continuum limit is a simple linear extrapolation in a^2 . The right, $s = 1$, point corresponds to 3 GeV

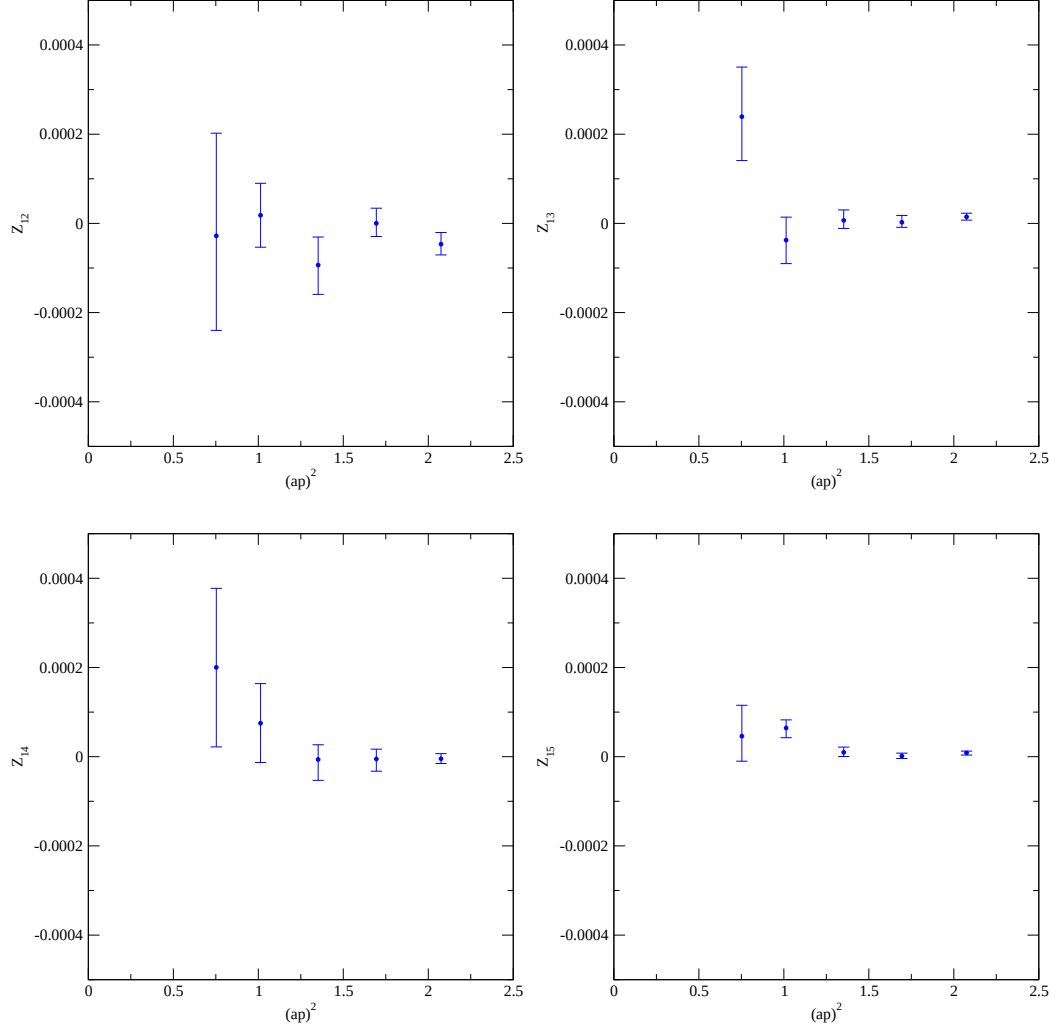


Figure 4.28: Mixing coefficient on 32^3 lattice for $O_1 = O_{VV+AA}$ and the operators $O_2 = O_{VV-AA}$, $O_3 = O_{SS-PP}$, $O_4 = O_{SS+PP}$ and $O_5 = O_{TT}$. The data shown has been extrapolated to the chiral limit.

Chapter 5

Eigenvectors of the Dirac Operator

In this chapter we calculate eigenvectors of the Dirac matrix. These can be used to accelerate the inversion of the Dirac matrix and we show results in section 5.4. We can also use them to decompose the matrix into its eigenvectors $|i\rangle$ and eigenvalues λ_i and approximate the propagator as,

$$S = D^{-1} \sim \sum_i \frac{|i\rangle\langle i|}{\lambda_i}, \quad (5.1)$$

truncating the sum after some finite number to give an estimate of the “all-to-all” propagator. How good is this estimate? The Dirac matrix is huge, $12V \times 12V$, where V is the number of lattice points and we have no hope of calculating a significant fraction of its eigenvectors for realistic lattices. However much of the important low energy physics in QCD is thought to be described by topological excitations of the vacuum. Classically they are instantons [118], non-trivial solutions to the classical field equations. These are thought to generate the QCD condensates, the ’t Hooft vertices [128] and they give important and non-perturbative contributions to physical observables. The index theorem, [11], connects the zero modes of the Dirac operator to the gluon action and suggests that the low lying eigenvectors couple to the low energy physics. Therefore the lowest few eigenvectors are the most important, with the high energy parts a small correction [107]. The accuracy of this truncation will be explored in detail in the next chapter, 6.

In the following we describe an algorithm for calculating the eigenvectors of large sparse matrices. We follow the description of the implicitly restarted Lanczos method in reference [138] closely and implement this approach. We then tune the algorithm for the Dirac matrices useful for lattice QCD with domain wall fermions. Then we demonstrate the effect of low mode deflation on the conjugate gradient algorithm. Finally we look at how the low modes couple to a classical instanton background and show that they are strongly spatially correlated with the instanton [57], [92].

5.1 Algorithms

For finding the eigenvalues of small dense matrices (where small means $O(100 \times 100)$) the implicitly shifted QR algorithm and its Hermitian simplification are preferred, [65], [66], [75], however these methods involve transformations of the matrix that alter its structure. We do not store the Dirac matrix as a large array but exploit its sparsity to write it as a routine that takes and returns a vector. For an $n \times n$ Dirac matrix this means multiplication is carried out in $O(n)$ operations but, if we destroyed the sparsity, doing a multiplication would take $O(n^2)$ operations. Also, we are typically satisfied with finding some small subset of the eigenvectors rather than all of them. What we would like is an algorithm that preserves the structure of the matrix while converging on a subset of wanted eigenvectors; this is accomplished by the Arnoldi (non-Hermitian) and Lanczos (Hermitian) algorithms. These algorithms are related to the QR algorithm [53] but do not change the structure of the matrix, so the $O(n)$ cost of multiplications is retained. We give a brief description of some of the mathematical ideas associated with the Arnoldi/Lanczos algorithms.

The j th Krylov subspace associated with matrix A and vector x is

$$\mathcal{K}_j(A, x) = \text{span}\{x, Ax, \dots, A^{j-1}x\}. \quad (5.2)$$

The power method shows us that repeatedly applying the operator A on the starting vector x brings the result closer and closer to the dominant eigendirection. The Krylov space then, is a subspace that comes close to spanning the space of the dominant eigenvectors.

An upper Hessenberg matrix is one of the form,

$$H = \begin{pmatrix} * & * & * & \dots & * & * \\ * & * & * & \dots & * & * \\ 0 & * & * & \dots & * & * \\ 0 & 0 & * & & * & * \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & * & * \end{pmatrix}. \quad (5.3)$$

Every matrix can be decomposed by a unitary transformation, V , into upper Hessenberg form [75],

$$H = V^\dagger A V. \quad (5.4)$$

The small matrix H is a similarity transform of A into the space spanned by the columns of V .

The Hessenberg decomposition and the Krylov space are related by the result:

- If $H = V^\dagger A V$ is upper Hessenberg and the columns of V are v_i then,

$$\text{span}\{v_1, v_2, \dots, v_j\} = \mathcal{K}_j(A, v_1). \quad (5.5)$$

The columns of V are an orthonormal basis for $\mathcal{K}_j(A, v_1)$ and so the eigenvectors of H will tell us about the dominant eigenvectors of A . To see this suppose we have an orthonormal basis for $\mathcal{K}_j(A, v_1)$,

$$\begin{aligned} \mathcal{K}_j(A, v_1) &= \text{span}\{v_1, Av_1, \dots, A^{j-1}v_1\} \\ &= \text{span}\{v_1, v_2, \dots, v_j\} \end{aligned} \quad (5.6)$$

where v_2 is constructed by orthogonalizing Av_1 against v_1 and v_3 is constructed by orthogonalizing A^2v_1 against v_1 and v_2 etc. The next Krylov space is,

$$\begin{aligned} \mathcal{K}_{j+1}(A, v_1) &= \text{span}\{v_1, Av_1, A^2v_1, \dots, A^jv_1\} \\ &= \text{span}\{v_1, (c_1v_1 + c_2v_2), A(c_1v_1 + c_2v_2), \dots, A^j(c_1v_1 + c_2v_2)\} \\ &= \text{span}\{v_1, v_2, Av_2, \dots, A^{j-1}v_2\} \\ &\vdots \\ &= \text{span}\{v_1, v_2, \dots, v_j, Av_j\}. \end{aligned} \quad (5.7)$$

This means an orthonormal basis for $\mathcal{K}_{j+1}$ can be obtained by orthogonalizing Av_j against $v_1, \dots, v_j$, instead of $A^j v_1$.

Now we rewrite the relation $AV = VH$ in terms of columns of V ,

$$Av_j = \sum_{i=1}^j v_i H_{ij} + v_{j+1} H_{j+1,j} \quad (5.8)$$

using the Hessenberg structure. This shows that v_{j+1} is a linear combination of $\{v_1, \dots, v_j, Av_j\}$ and, using $v_i^\dagger Av_j = H_{ij}$ and the result from the previous paragraph, the result that the columns of V are an orthonormal basis for the Krylov space follows by induction.

5.1.1 Arnoldi Algorithm

The Arnoldi algorithm uses these ideas to explicitly construct the matrices H and V . Given an orthonormal basis for $\mathcal{K}_j(A, x) = \text{span}\{v_1, \dots, v_j\}$ we want to construct an orthonormal basis for the next Krylov space

$$\begin{aligned} \mathcal{K}_{j+1}(A, x) &= \text{span}\{x, Ax, \dots, A^j x\} \\ &= \text{span}\{v_1, \dots, v_j, Av_j\}. \end{aligned} \quad (5.9)$$

Let

$$f_j = Av_j - \sum_{i=1}^j v_i (v_i^\dagger Av_j) \quad (5.10)$$

where f_j is constructed to be orthogonal to $\text{span}\{v_1, \dots, v_j\}$ and

$$v_{j+1} = \frac{f_j}{|f_j|}. \quad (5.11)$$

By construction, $v_i^\dagger v_{j+1} = 0$ for $i = 1, 2, \dots, j$ and therefore

$$\mathcal{K}_{j+1}(A, x) = \text{span}\{v_1, \dots, v_j, v_{j+1}\}, \quad (5.12)$$

so the vectors v_i are the columns of a unitary transformation of A into an upper Hessenberg matrix, H given by,

$$H_{ij} = v_i^\dagger Av_j. \quad (5.13)$$

To see explicitly that H is upper Hessenberg we can decompose Av_k in terms of basis vectors of $\mathcal{K}_{k+1}(A, x)$

$$Av_k = \sum_{i=1}^{k+1} c_i v_i. \quad (5.14)$$

For $k \leq j+2$

$$H_{jk} = v_j^\dagger Av_k = v_j^\dagger \left(\sum_i^{k+1} c_i v_i \right) = 0, \quad (5.15)$$

this means the matrix H is upper Hessenberg and we have the decomposition,

$$H_K = V_K^\dagger AV_K. \quad (5.16)$$

The Arnoldi iteration [134] produces an upper Hessenberg matrix H_K and K vectors, $V_K = (v_1, \dots, v_K)$ such that,

$$\begin{aligned} AV_K &= V_K H_K + f_K e_K^\dagger \\ V_K^\dagger f_K &= 0 \end{aligned} \quad (5.17)$$

so that equation 5.16 holds. The vector e_K has K elements and $e_K^\dagger = (0, 0, \dots, 1)$. We have constructed a similarity transformation on A that gives a partial Hessenberg decomposition in the subspace $\text{span}\{v_1, \dots, v_K\} = \mathcal{K}_K(A, v_1)$.

Algorithm 1 Arnoldi algorithm

Require: $v \neq 0$

```

 $v_1 = \frac{v}{|v|}$             $V_1 = [v_1]$ 
 $w = Av_1$             $h = v_1^\dagger w$ 
 $f_1 = w - v_1 h$       $H_1 = [h]$ 
for  $j = 1, 2, \dots, K-1$  do
   $\beta_j = |f_j|$ 
   $v_{j+1} = \frac{f_j}{\beta_j}$ 
   $V_{j+1} = [V_j, v_{j+1}]$ 
   $\tilde{H}_j = \begin{bmatrix} H_j \\ \beta_j e_j^\dagger \end{bmatrix}$ 
   $w = Av_{j+1}$ 
   $h = V_{j+1}^\dagger w$ 
   $f_{j+1} = w - V_{j+1} h$ 
   $H_{j+1} = \begin{bmatrix} \tilde{H}_j \\ h \end{bmatrix}$ 
end for
```

The matrix H_K is small ($K \times K$) and the implicit QR algorithm for dense matrices can be applied to find the Ritz eigenvalues θ_i and eigenvectors $y_i = V_K s_i$, with

$$H_K s_i = \theta_i s_i. \quad (5.18)$$

If V_K consisted of exact eigenvectors of A then H_K would be diagonal. The Krylov space spans the most dominant eigenvectors of A therefore we expect that some subset of the Ritz values will be good approximations to the true eigenvalues and eigenvectors of A . To check how good we may compute,

$$(A - \theta_i)y_i = AV_K s_i - V_K \theta_i s_i = (AV_K - V_K H_K)s_i = f_K e_K^\dagger s_i \quad (5.19)$$

hence the magnitude of the residual is equal to,

$$|f_K|(s_i)_K = H_{K+1K}(s_i)_K, \quad (5.20)$$

the last component of s_i times H_{K+1K} , which is positive. This is useful in deriving a stopping procedure.

5.1.2 Lanczos Algorithm

If the matrix is Hermitian the Arnoldi algorithm can be simplified. First notice that $V_K^\dagger A V_K$ is Hermitian if A is. This is just the upper Hessenberg matrix H_K constructed in the Arnoldi process, hence H_K is Hermitian and so must be a tridiagonal matrix T_K ,

$$V_K^\dagger A V_K = T_K \quad (5.21)$$

$$T_K = \begin{pmatrix} \alpha_1 & \beta_1 & 0 & 0 & 0 \\ \beta_1 & \alpha_2 & \beta_2 & 0 & 0 \\ 0 & \beta_2 & \alpha_3 & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & \beta_{K-1} \\ 0 & 0 & 0 & \beta_{K-1} & \alpha_K \end{pmatrix}. \quad (5.22)$$

This means that in the orthogonalization step,

$$f_j = w - V_j h \quad (5.23)$$

$$f_j = Av_j - \sum_{i=1}^j v_i v_i^\dagger Av_j,$$

the sum only has two nonzero contributions, since $v_i^\dagger Av_j = T_{ij}$ and the residual can be calculated from the three term recurrence,

$$f_j = Av_j - \alpha_j v_j - \beta_{j-1} v_{j-1}. \quad (5.24)$$

Because of the mutual orthogonality of all the v_i we have

$$v_{j+1}^\dagger f_j = v_{j+1}^\dagger Av_j = \beta_j \quad (5.25)$$

and

$$v_j^\dagger (Av_j - \beta_{j-1} v_{j-1}) = v_j^\dagger Av_j = \alpha_j. \quad (5.26)$$

The Arnoldi algorithm with these modifications becomes the Lanczos algorithm. The recurrence above means that only the last residual needs to be kept, we call the remaining one $f_K = r$.

Algorithm 2 Lanczos algorithm

Require: $q \neq 0$

$v_1 = \frac{v}{ v }$	$V_1 = [v_1]$
$r = Av_1$	$\alpha_1 = v_1^\dagger r$
$r = r - \alpha_1 v_1$	$\beta_1 = r $

for $j = 2, 3, \dots, K$ **do**

$v_j = \frac{r}{\beta_{j-1}}$	
$V_j = [V_{j-1}, v_j]$	
$r = Av_j - \beta_{j-1} v_{j-1}$	
$\alpha_j = v_j^\dagger r$	
$r = r - \alpha_j v_j$	
$\beta_j = r $	

end for

The convergence is monitored again by the product of the last component of the Ritz eigenvector times β_K . In the Hermitian case more can be said about the convergence using the Kaniel-Paige-Saad bounds, see [113], [87] and

references therein. These say we get faster convergence for eigendirections strongly represented in the initial vector, for eigenvalues on the ends of the spectrum and that the better separated the eigenvalues are from each other the faster they converge. For the Dirac matrix this is ideal since we are only interested in the lowest eigenvalues and these tend to be well separated from each other compared to the dense high modes.

5.2 Implementation

5.2.1 Implicit Restarting

In practical implementations we run the Arnoldi or Lanczos algorithms for a fixed number of steps, compute the Ritz vectors and use these to restart the process. This is continued until the residuals are small enough. The new starting vector is usually constructed as some linear combination of the Ritz vectors, since, in the limit when these are exact, the residual will vanish. This is known as explicit restarting. Instead of directly constructing the new start vector there is an alternative known as implicit restarting, [126], [35].

Consider the simple case where our starting vector is a linear combination of two eigenvectors,

$$v = c_1 v_1 + c_2 v_2. \quad (5.27)$$

If we knew the eigenvalue associated with v_1 , λ_1 , we could restart with

$$(A - \lambda_1)v = c_1(\lambda_1 - \lambda_1)v_1 + c_2(\lambda_2 - \lambda_1)v_2 \quad (5.28)$$

which removes the component in the known v_1 direction. So by applying a polynomial

$$p(A) = (A - s_1) \dots (A - s_n) \quad (5.29)$$

with shifts s_n to the starting vector we can remove or reduce components in some eigendirections and converge to the desired subspace. If we have a truncated QR decomposition we can calculate the spectrum of H_M , $\sigma(H_M) = \{\theta_1, \dots, \theta_M\}$ and divide it into K wanted and P unwanted components. In our case this will mean taking the P largest or smallest eigenvalues as shifts. We do not apply $p(A)$ directly but implicitly, through a QR transformation, as shown in algorithm

3, where H is a tridiagonal or Hessenberg matrix depending on whether A is Hermitian or not.

Algorithm 3 Implicitly Restarted Arnoldi/Lanczos

Require: $M > K$
 $P = M - K$

 Compute the factorization $AV_M = V_M H_M + f_M e_M^\dagger$
repeat
 $Q = I$
for $i = 1, \dots, P$ **do**
 $Q_i R_i = H_M - \theta_i I$
 $Q = Q Q_i$
 $H_M = Q_i^\dagger H_M Q_i$
end for
 $\beta_K = H_M(K+1, K)$
 $\sigma_K = Q(M, K)$
if Arnoldi **then**
 $f_K = v_{K+1} \beta_K + f_M \sigma_K$
end if
if Lanczos **then**
 $r = v_{K+1} \beta_K + r \sigma_K$
end if
 $V_K = V_M(1:M)Q(1:M, 1:K)$
 $H_K = H_M(1:K, 1:K)$
 $\rightarrow AV_K = V_K H_K + f_K e_K^\dagger$

 Extend to an $M = K + P$ step factorization $AV_M = V_M H_M + f_M e_M^\dagger$
until convergence

To see how the implicit restart works let us go through it. After M steps of Arnoldi/Lanczos the QR decomposition,

$$QR = H_M - \theta I \quad (5.30)$$

is constructed for a given shift θ . Using the fact that Q is orthogonal

$$\begin{aligned} AV_M Q &= V_M Q Q^\dagger H_M Q + f_M e_M^\dagger Q \\ AV'_M &= V'_M H'_M + f_M e_M^\dagger Q \end{aligned} \quad (5.31)$$

with $V'_M = V_M Q$. It can be shown [35] that

$$e_M^\dagger Q = \tilde{e} = (0, \dots, 0, \tilde{e}_K, \dots, \tilde{e}_M), \quad (5.32)$$

the first non-zero row of Q is the K^{th} one. By equating the first K columns above we obtain,

$$AV_K = V_K H_K + f_K e_K \quad (5.33)$$

which is a K step Arnoldi factorization. This is the factorization we would have obtained had we started with the vector $p(A)v$. To show this,

$$\begin{aligned} AV_M &= V_M H_M + f_M e_M^\dagger \\ AV_M &= V_M (QR + \theta I) + f_M e_M^\dagger \\ (A - \theta I)V_M &= V'_M R + f_M e_M^\dagger \end{aligned} \quad (5.34)$$

multiplying by e_1 gives the first column,

$$(A - \theta I)V_M e_1 = V'_M R e_1. \quad (5.35)$$

Since R is upper triangular

$$R e_1 = (\rho, 0, 0, \dots, 0) \quad (5.36)$$

thus,

$$v'_1 = \frac{1}{\rho} (A - \theta I)v_1. \quad (5.37)$$

The new factorization is equivalent to one started with $(A - \theta I)v_1$ instead of v_1 and as we repeat the process with the other shifts we effectively multiply by the other factors of $p(A)$.

If the initial vector is written in terms of eigenvectors of A , q_i , then

$$\begin{aligned} v &= \sum_i c_i q_i \\ p(A)v &= \sum_i c_i p(\lambda_i) q_i. \end{aligned} \quad (5.38)$$

The components for which $p(\lambda_i)$ is small, the unwanted eigenvalues, will be reduced. Typically for Dirac matrices the spectrum gets denser for eigenvalues of larger magnitude and when we have converged on $j \sim K$ eigenvalues it may be that the unwanted and wanted eigenvalues are close together, so that using the approximations to the unwanted eigenvalues as shifts makes convergence to the

wanted values slower. As the number of converged j approaches K we reduce P , however not so much that we fail to shift out eigenvalues that really are unwanted. The empirical rule of thumb $K - j < 4 \rightarrow P = P - 1$ is found to work well and reduces the number of Arnoldi/Lanczos iterations in some cases, though this will rarely be important in practical applications.

5.2.2 Locking

Once we have converged on an eigenvalue, signified by a small last component of the Ritz vector, it is convenient to lock the eigenpair (x, θ) and carry on the iteration without disturbing the converged pair; we achieve this using the method of [95]. Begin with

$$H_K y = y\theta \quad (5.39)$$

where $|y| = 1$ and the last component $|y_K| = \eta < \epsilon|H|$ where ϵ is some small number. We construct Q as in reference [95] such that

$$\begin{aligned} Qe_1 &= y \\ e_K^T Q &= (\eta, \tau e_{k+1}^T) \\ \eta^2 + \tau^2 &= 1. \end{aligned} \quad (5.40)$$

Then we apply the transformation,

$$\begin{aligned} AV &= VH + f e_K^T \\ \rightarrow AVQ &= VQQ^\dagger H Q + f e_K^T Q \\ A[v_1 V_2] &= [v_1 V_2] \begin{bmatrix} \theta & h^T \\ 0 & H_2 \end{bmatrix} + f(\eta, \tau e_{k+1}^T) \end{aligned} \quad (5.41)$$

the matrix H_2 is brought into upper Hessenberg form again and then subsequent restarting and Lanczos steps take place only on the lower block.

5.2.3 Gram Schmidt

The Arnoldi and Lanczos algorithms use Gram Schmidt orthogonalization. Gram Schmidt is known to suffer from large rounding errors which slows down convergence of the Arnoldi/Lanczos. The Arnoldi algorithm explicitly orthogonalizes

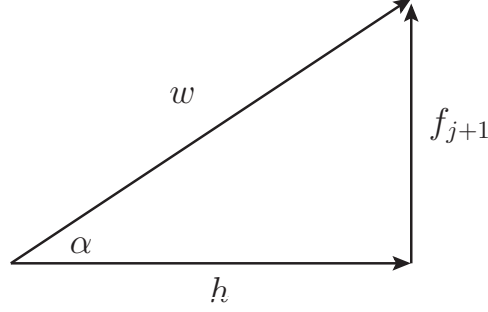


Figure 5.1: Showing the vectors w , h and f_j schematically in two dimensions.

against the previous v 's but the Lanczos only explicitly orthogonalizes against the previous two, which means it suffers more from rounding errors. We use an iterative refinement procedure that ought to ensure the vectors produced by the Gram Schmidt process are numerically orthogonal. In the Arnoldi algorithm we have the vectors, w , $h = V_{j+1}^\dagger w$ and $f_{j+1} = (1 - V_{j+1}V_{j+1}^\dagger)w$. The vector h is the projection of w into the space spanned by the vectors of V_{j+1} and f is its complement, see figure 5.1. Thus the angle α in

$$\tan(\alpha) = \frac{|f_{j+1}|}{|h|} \quad (5.42)$$

is a measure of how close w is to the space already constructed, the closer it is the worse we expect the roundoff error to be. If it is small we can subtract off the components of f in the space spanned by V_{j+1} to enforce orthogonality.

Algorithm 4 Iterative Refinement

```

if  $\frac{|f_j|}{|h|} < \eta$  then
   $s = V_j^\dagger f_j$ 
   $f_j = f_j - V_j s$ 
   $h = h + s$ 
end if

```

The parameter $\eta = \tan(\alpha)$ in the algorithm is usually set to 1 and for the Lanczos algorithm $|h| = \sqrt{\alpha_j^2 + \beta_{j-1}^2}$. Lanczos may need several steps of iterative refinement, after the first step the test is $\frac{|f_{j+1}|}{|s|} < \eta$. If after a few refinements an orthogonal vector has not been found we terminate the process and check for an invariant subspace.

5.2.4 Chebyshev filters

The matrix whose eigensystem is to be found may be replaced with a polynomial of the matrix $p(A)$ which has better spectral properties, namely that the wanted eigenvalues are large and well separated. The largest eigenvalues of $p(A)$,

$$p(A)v_i = p(\lambda_i)v_i \quad (5.43)$$

can be found in fewer iterations than the lowest eigenvalues of A . Since the eigenvectors of A and $p(A)$ are common the eigenvalues of A can be obtained from the Rayleigh quotient,

$$\lambda_i = \frac{v_i^\dagger A v_i}{v_i^\dagger v_i}, \quad (5.44)$$

once the eigenvectors have been found. The polynomial here is applied explicitly and will be more expensive than doing a single multiplication by A , thus there is a trade off between a smaller number of iterations and a larger cost for each one. We will find in favour of high order polynomials for some matrices in section 5.3.1 and no polynomial at all for others. We concentrate here on filters for the Lanczos algorithm [116], however filters for the Arnoldi algorithm have also been proposed [19].

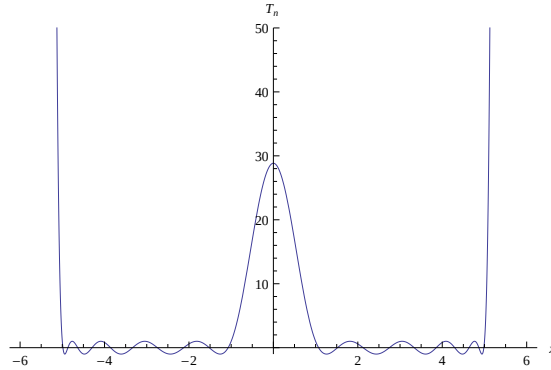


Figure 5.2: The polynomial $T_{10}(q(x; \alpha = 5, \beta = 1))$.

In order to compute the lowest eigenvalues of A we need a polynomial that is nearly zero for the high modes and rapidly increasing for the low ones, separating

them out from each other. First make the transformation,

$$q(A; \alpha, \beta) = \frac{2A^2 - (\alpha^2 + \beta^2)}{(\alpha^2 - \beta^2)} \quad (5.45)$$

with α and β known constants to be specified later. Now construct the n^{th} order Chebyshev polynomial, $T_n(q(A))$, by the recurrence ,

$$T_n(q(A)) = 2q(A)T_{n-1}(q(A)) - T_{n-2}(q(A)). \quad (5.46)$$

This polynomial is plotted for particular values of n , α and β in figure 5.2. $q(x)$ has the property that between $x = \beta$ and $x = \alpha$ it goes from -1 to 1 . The Chebyshev polynomials are small in the interval $[-1, 1]$ and rapidly rising outside, thus $T_n(q(x))$ will be large for $x > \alpha$ and $x < \beta$. If we choose α as the largest eigenvalue of the matrix A (or larger) and β and the largest eigenvalue we want to compute we will greatly enhance the wanted part of the spectrum and damp the part in the range $[\beta, \alpha]$.

To get interior eigenvalues we can apply a shift μ

$$q(A; \alpha, \beta, \mu) = \frac{2(A - \mu)(A - \mu) - (\alpha^2 + \beta^2)}{(\alpha^2 - \beta^2)} \quad (5.47)$$

for Hermitian A . Now μ controls the peak position. We put $\alpha = \alpha + \mu$, the maximum eigenvalue plus the shift, this is necessary if A has negative eigenvalues, as is the case for the Hermitian Dirac operators H , and β controls the width of the peak. Unfortunately narrow peaks will also be very small and will not enhance the nearby spectrum as much, one has to strike a balance between the width and the size of the peak see figure 5.3.

5.3 Algorithm Tests

The Arnoldi algorithm can find eigenvalues and eigenvectors of non-Hermitian systems, however, by restricting it to Hermitian systems and using the Lanczos algorithm we get a substantial speed up. We can also use Chebyshev acceleration, making Lanczos even faster. Furthermore we intend to use the eigenvectors for accelerating CG solves which only requires eigenvectors of the manifestly Hermitian matrix $D^\dagger D$, section 3.2. Low mode averaging, chapter 6, also only

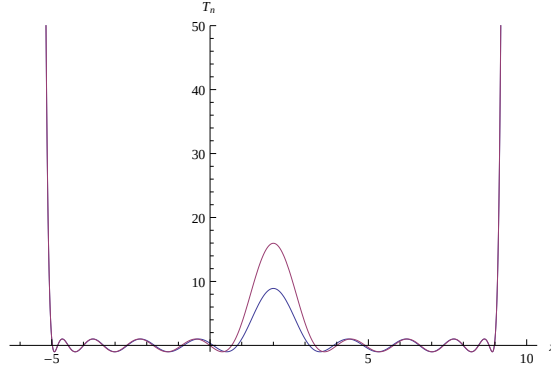


Figure 5.3: The polynomials $T_{10}(q(x; \alpha = 5, \beta = 1, \mu = 2))$ (blue, smaller peak) and $T_{10}(q(x; \alpha = 5, \beta = 1.2, \mu = 2))$ (red, larger peak).

uses eigenvectors of Hermitian matrices. Using eigenvectors of non-Hermitian matrices for low mode averaging has been attempted [13] without success. Reference [13] also found that generating the eigenvectors was much more expensive (which we have also found). We leave the Arnoldi algorithm here and consider only the Lanczos algorithm for Hermitian systems from now on.

From appendix A we know the eigenvalues of the Hermitian overlap operator H_{ov} exactly in the unit gauge thus we can calculate the same with our eigensolver as a test. We choose a small lattice, 4^4 , with $L_s = 4$ and calculate all of the eigenvalues. The unit gauge spectrum has many degenerate eigenvalues which are correctly found by our eigensolver, see figure 5.4. The unit gauge is awkward and not a very realistic case due to this large degeneracy which is lifted in a real gauge field background. For tests we use a small, $4^3 \times 8$, $L_s = 4$, lattice generated with the Iwasaki gauge action, $\beta = 2.13$, $M_5 = 1.8$, with light quark masses $m_l = 0.04$ and the strange sea quark mass, $m_s = 0.04$. We use a valence quark of mass $m = 0.04$ in our tests unless otherwise specified. The trajectories were generated using the RHMC algorithm and we use the configurations at molecular dynamics time 400 and 450 of the molecular dynamics trajectory. We will use three Dirac operators these are: the effective overlap operator $H_{ov} = \gamma_5 D_{ov}$, equation 1.99; the Hermitian domain wall operator, $H_{DWF} = \gamma_5 \mathcal{R}_5 D_{DWF}$, equation 1.49; and the preconditioned domain wall operator $D_{oo}^\dagger D_{oo}$, equation 3.33. Evaluating the operator H_{ov} requires the inversion of the Pauli-Villars matrix, see section 1.7. We do this inversion to an accuracy of $\epsilon/10$, where ϵ is the desired residual of the eigenvectors. Doing it to greater accuracy does not change the residual at the

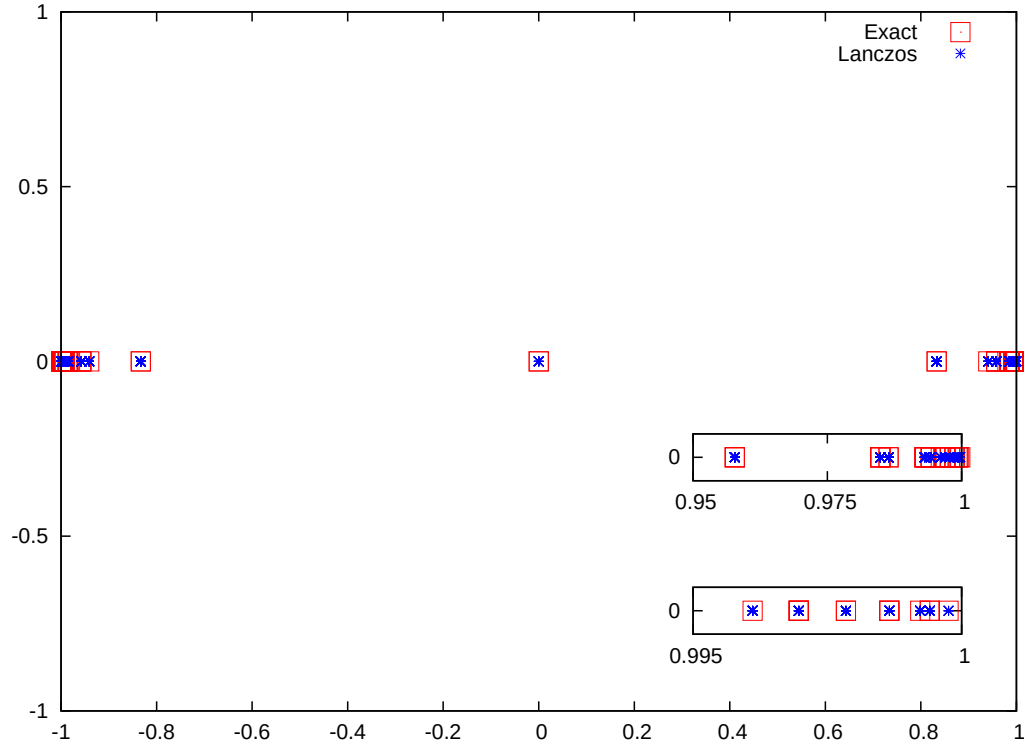


Figure 5.4: We show the exact result of appendix A equation A.24 (red squares) compared with the result of our Lanczos eigensolver (blue stars) for H_{ov} in the unit gauge after a random gauge transformation, the eigenvalues are gauge invariant and exactly match the analytical ones.

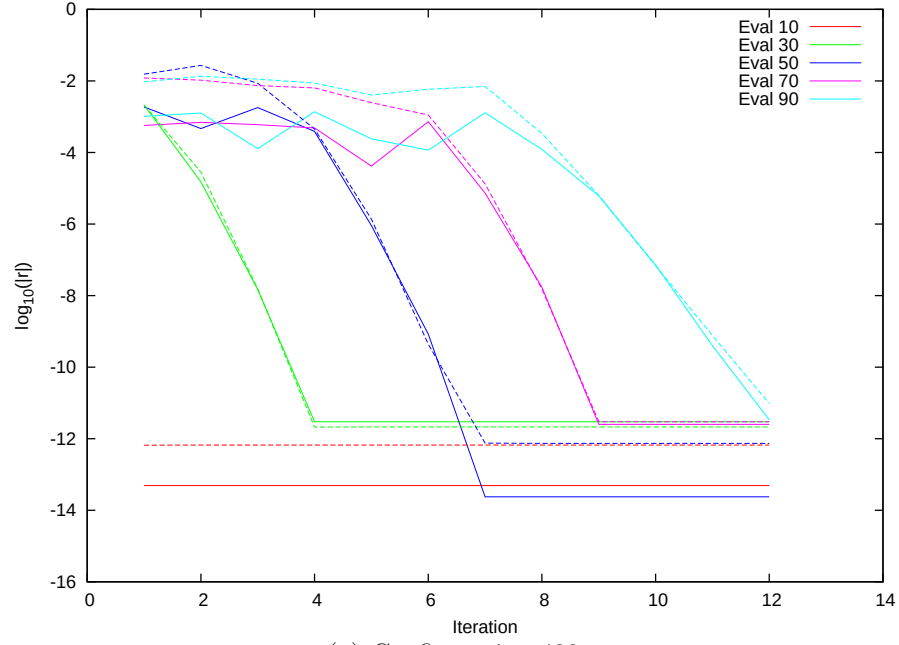
required level of precision and only makes each multiplication more expensive.

We first plot the estimated and true residual against the iteration number for eigenvectors of H_{ov} on the 400 and 450 configuration in figure 5.5. We see that the estimate of the residual from equation 5.20 (solid lines) tends to estimate the true residual (dashed lines) quite well and the true residual of final eigenvectors is at the level of 10^{-11} , the accuracy of the inner CG. Using the estimate of the residual as a stopping criteria is good enough. We also find that using Chebyshev acceleration gives us very small residual estimates in very few iterations see figure 5.6. The true residuals are of the order 10^{-14} which is the accuracy to which we solve the $HS = \theta s$ eigensystem. Thus where Chebyshev acceleration works, which turns out to be for the 5d operators H_{DWF} and $D_{oo}^\dagger D_{oo}$, we can demand very small residuals and obtain very precise eigenvectors.

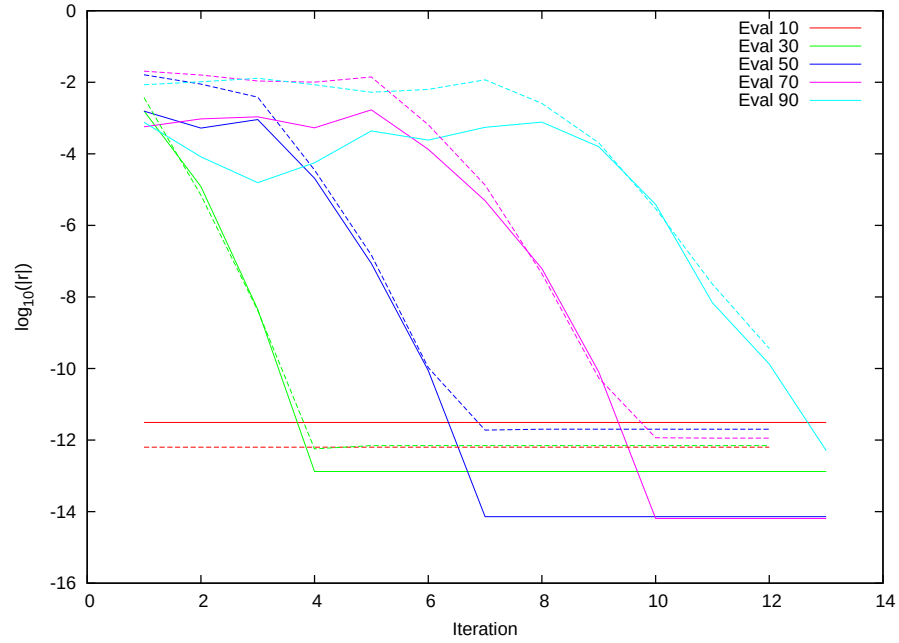
5.3.1 Tuning

Chebyshev Parameters

There are several parameters to tune in the Lanczos algorithm. First we consider the parameters α and β of the Chebyshev polynomial, equation 5.47. The eigenvalues in this section will be locked when the residual is smaller than $\epsilon = 10^{-10}$. Our problem is to find the eigenvectors of the Dirac matrix on different gauge field backgrounds. Different gauge fields mean that the eigenvalues will be different, however most of the detailed structure of the gauge fields is coupled to the low modes, the higher modes are affected less and less by topological excitations and gauge field lumps and thus the higher eigenvalues vary less. The spectral radius, λ_R , can be quickly calculated by using a Chebyshev polynomial with the offset, β , set to zero and $\alpha < \lambda_R$. The Lanczos process then rapidly converges to the highest eigenvalue, λ_R , which hardly changes from configuration to configuration. We set $\alpha = \lambda_R + \eta_\alpha$ with η_α a small positive number to ensure that $\alpha \geq \lambda_R$ for all configurations. If we calculate K eigenvalues on one configuration using a rough estimate for β we get λ_K , the highest wanted eigenvalue, on that configuration. Setting $\beta = \lambda_K + \eta_\beta$ with η_β a small positive number ensures $\beta \geq \lambda_K$ for all configurations. Generally for very small numbers of eigenvalues, $O(10)$, η_β should be around 10% of β to be safe, as the number required increases η_β can be reduced.

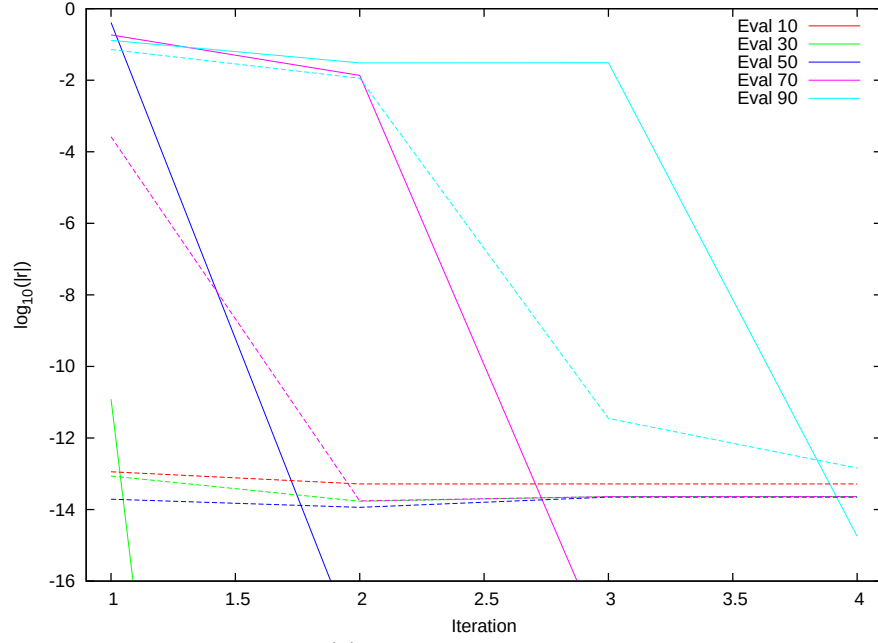


(a) Configuration 400

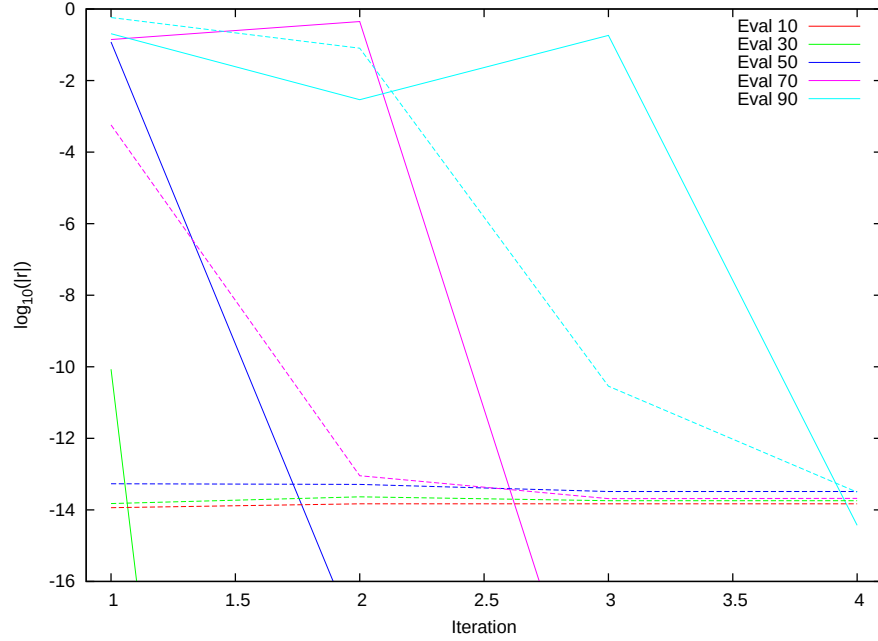


(b) Configuration 450

Figure 5.5: The estimate of the residual equation 5.20 (solid lines) and the true residual $|Hv - \lambda v|$ (dashed lines) for the operator H_{ov} on the small lattices described in the text. When the residual estimate is less than 10^{-10} we lock the eigenvector as in section 5.2.2 and its value no longer changes. We show the residuals for the 10th, 30th, 50th, 70th and 90th eigenvectors.



(a) Configuration 400



(b) Configuration 450

Figure 5.6: The estimate of the residual equation 5.20 (solid lines) and the true residual $|D_{oo}^\dagger D_{oo} v - \lambda v|$ (dashed lines) for the operator $D_{oo}^\dagger D_{oo}$ on the small lattices described in the text. When the residual estimate is less than 10^{-10} we lock the eigenvector as in section 5.2.2 and its value no longer changes. We show the residuals for the 10th, 30th, 50th, 70th and 90th eigenvectors. The residual estimate on convergence is much less than the true residual, which is limited by the Ritz QR solve.

This process is illustrated in table 5.1. We pick a Chebyshev polynomial of order $n = 40$. M , K and n will be tuned later, for now we pick reasonable values. The cost is measured in terms of the number of matrix vector multiplications, N_{DWF} , and number of restarts, N_{re} , there is additional overhead involved in solving $K \times K$ subsystems, QR decomposition for the implicit restart, orthogonalization etc. but most of the time is spent doing multiplications. We search for eigenvectors of H_{DWF} . Starting at the top of the table we take configuration number 400, calculate its spectral radius and find $\lambda_R = 8.2064974838$, we take a comfortable margin of error and set $\alpha = 8.3$. We then guess that $\lambda_K < \beta = 2.0$. We find the first hundred eigenvalues and the true value of $\lambda_K = -1.0273810227$, we take another comfortable margin of error and $\beta = 1.1$. Using these parameters on configuration 450 the Lanczos process converges in three to four times fewer iterations. The same α and β can be used on every configuration. Also, even though the eigenvalues will increase or decrease when increasing or decreasing the valence mass, the change is usually small enough that α and β do not need to be recalculated.

Configuration 400						
M	K	n	α	β	N_{DWF}	N_{re}
20	1	40	5	0	3121	1
120	100	40	8.3	2.0	27300	11
Configuration 450						
120	100	40	8.3	1.1	14500	3

Table 5.1: Tuning α and β on one configuration greatly reduces the cost of the eigensolver on the other configurations.

Tuning M , n

Now we look at tuning M and n for each of the operators H_{ov} , H_{DWF} and $D_{oo}^\dagger D_{oo}$. To do this we look at the total time as a measure of the cost and plot this as a function of M for different values of n . H_{ov} requires one multiplication by D_{DWF} and solving the equation, $D_{DWF}^{PV}x = b$, which is achieved as in section 1.7. In order to use the same cost measure we count L_s Wilson multiplications the same as one D_{DWF} . Now we use realistic configurations to tune the eigensolver. These are $16^3 \times 32$, $L_s = 16$, $M_5 = 1.8$, Iwasaki gauge $\beta = 2.13$ configurations with

$m_s = 0.04$ as reported in [1]. There are three ensembles with light quark masses $m = \{0.01, 0.02, 0.03\}$ and we set the valence quark mass (the m in D_{DWF}) equal to the light sea quark mass. These will be referred to as the “16³” ensembles. For the tests below we take configuration number 3000 of the $m = 0.01$ ensemble and tune the Chebyshev parameters as in the previous section.

Figures 5.3.1 (a), (b) and (c) show the effect of varying the order of the Chebyshev polynomial for different numbers of unwanted vectors in the Lanczos process. We have two competing requirements, the cost of a multiplication and the cost of the linear algebra. The three term recurrence in the Lanczos process is not very good at maintaining orthogonality amongst all of the M Lanczos vectors and iterative refinement must be done quite often. Also after every iteration of the Lanczos process we have to solve an $M \times M$ system which can become costly for large enough M . The optimal values of M and n will depend on the operator, though we have some expectations. The 4d operator D_{ov} is very expensive to apply (requiring an inner CG solve) so the optimal parameters for this will minimize the number of multiplications. For the 5d cases the domain wall operator has been coded very efficiently [29] and the cost of inner products goes up by a factor L_s , thus the optimal parameters in this case are such that trading more expensive multiplications for fewer restarts and less linear algebra reduces the total cost.

Figure 5.3.1 implies that for $D_{oo}^\dagger D_{oo}$ one should choose very few unwanted Lanczos vectors and a Chebyshev order of $O(80)$ to converge in minimal total time. We choose $M = 120$ because we observe that for very small numbers of unwanted eigenvectors, $O(10)$, the estimated and true residuals differ more significantly, leading to underestimates of the true residual and premature convergence. Figure 5.3.1 shows that as the cost of linear algebra increases (we have to work with full parity vectors instead of half parity ones) the total time goes up accordingly and is minimized by taking an even higher order Chebyshev polynomial, $O(140)$. The high order Chebyshev polynomials are what makes this algorithm effective. For the 5d unpreconditioned case in 24 hours of machine time, equivalent to computing 72 propagators on the same lattice, we are able to converge only 75 eigenvectors of H_{DWF} to the same accuracy without using Chebyshev acceleration. For the preconditioned system without Chebyshev acceleration 100 eigenvectors can be found in a time equivalent to $N_{prop} = 7.5$.

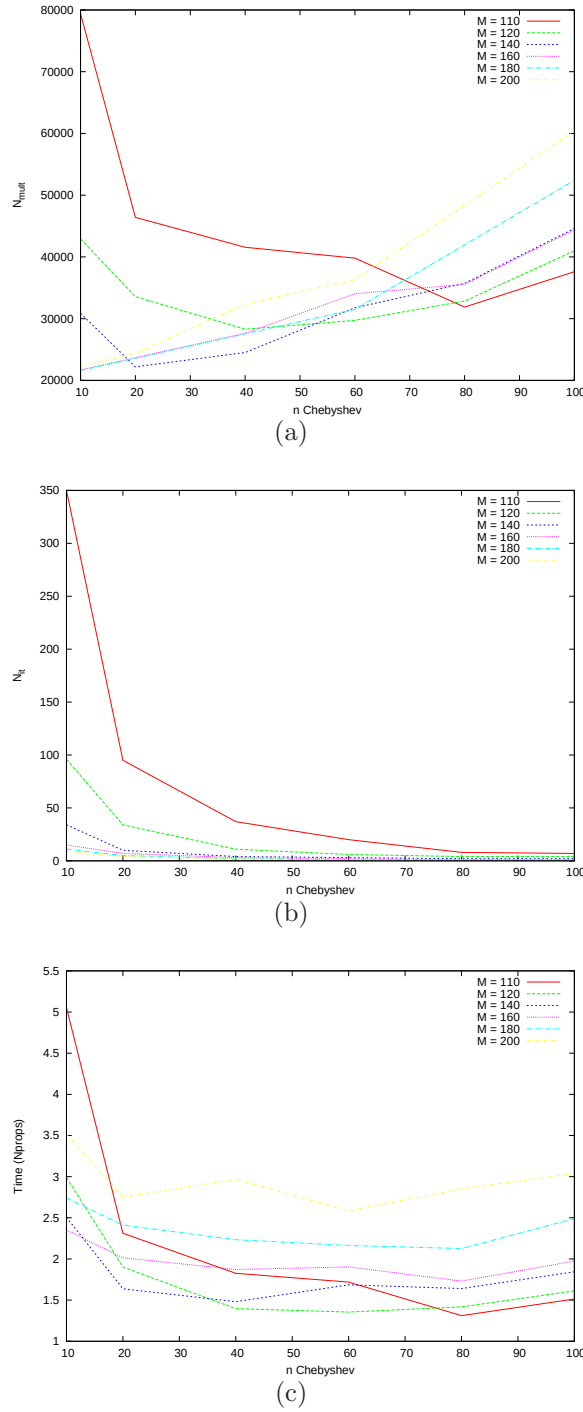
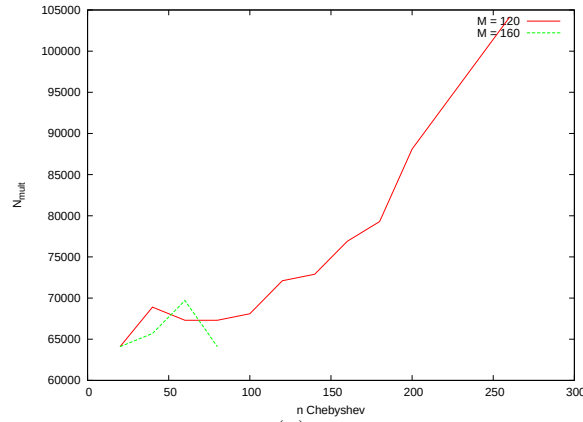
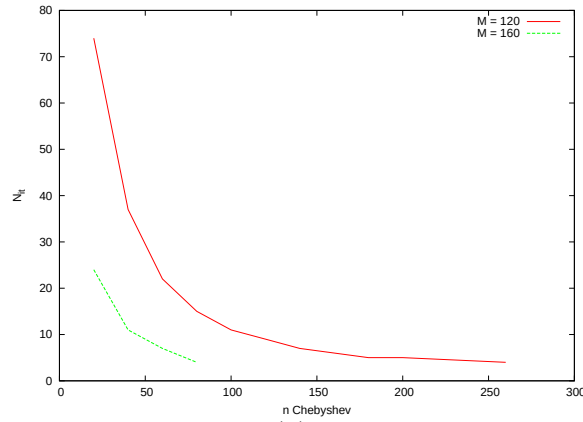


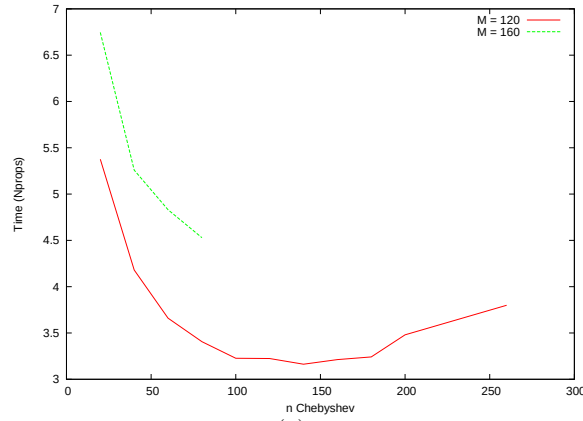
Figure 5.7: Tuning M and n for $D_{oo}^\dagger D_{oo}$, $K = 100$ wanted eigenvectors and $M - K$ unwanted. (a) Number of multiplications vs. polynomial order n for different M 's. Increasing the polynomial order increases N_{mult} approximately linearly. (b) Number of restarts vs. n for different values of M . Higher order Chebyshev leads to far fewer restarts. (c) Total time as a fraction of the time to compute a point source propagator on the same machine vs. n for different M . The optimal choice is $M = 120$, $n = 80$.



(a)



(b)



(c)

Figure 5.8: Tuning M and n for H_{DWF} , $K = 100$ wanted eigenvectors and $M - K$ unwanted. (a) Number of multiplications vs. polynomial order n for different M 's. (b) Number of restarts vs. n for different values of M . Higher order Chebyshev leads to far fewer restarts. (c) Total time as a fraction of the time to compute a point source propagator on the same machine vs. n for different M . The optimal choice is $M = 120$ and $n = 140$.

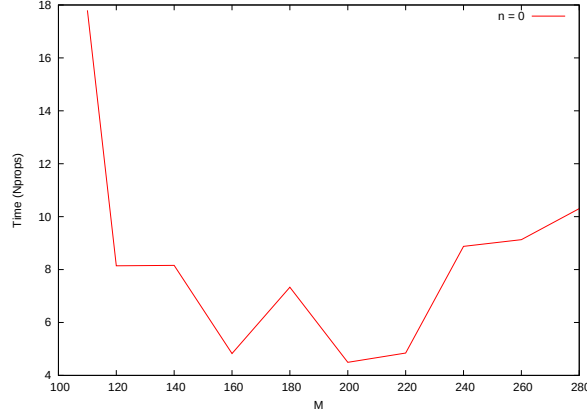


Figure 5.9: Tuning M for H_{ov} , $K = 100$ wanted eigenvectors and $M - K$ unwanted. Total time as a fraction of the time to compute a point source propagator on the same machine. The optimal choice is to have M around 200. Using a Chebyshev polynomial of order higher than zero is much slower.

Without Chebyshev the optimal M is 160. Compared to our optimal Chebyshev parameters, which give $N_{prop} = 1.5$ for $D_{oo}^\dagger D_{oo}$, this is a $5\times$ speed up. For the unpreconditioned calculation Chebyshev filtering makes the calculation possible at reasonable expense and for the preconditioned operator gives a factor 5 speed up.

Figure 5.3.1 shows that, for the four dimensional overlap operator, one hundred eigenvectors can be computed in the same time as about four propagator inversions and the number of wanted and unwanted eigenvectors should be about equal. In this figure we have set the residual to be less than 10^{-4} and the CG tolerance for the Pauli-Villars inverse to be 10^{-5} . As we will see later, low precision is acceptable for low mode averaging, chapter 6.

5.4 Deflation

Given the N lowest of the M possible eigenvectors of $D_{oo}^\dagger D_{oo}$,

$$D_{oo}^\dagger D_{oo} |i\rangle = \lambda_i |i\rangle, \quad (5.48)$$

we show how to use them to speed up CG convergence. The eigenvectors are orthogonal, since the matrix is Hermitian and assumed normalized to one and

the eigenvalues are real. For realistic gauge configurations the lowest eigenvalues are relatively sparse on the real line, so if we could remove the lowest N we would improve the condition number and have a denser spectrum, both good for convergence of CG.

To do this first observe,

$$D_{oo}^{-1} = (D_{oo}^\dagger D_{oo})^{-1} D_{oo}^\dagger = \sum_{i=0}^M \frac{|i\rangle\langle i|}{\lambda_i} D_{oo}^\dagger. \quad (5.49)$$

Then rewrite $(D_{oo}^\dagger D_{oo})^{-1}$ in terms of low and high mode parts,

$$\begin{aligned} (D_{oo}^\dagger D_{oo})^{-1} &= \sum_{i=0}^M \frac{|i\rangle\langle i|}{\lambda_i} \\ &= \sum_{i=0}^N \frac{|i\rangle\langle i|}{\lambda_i} + \sum_{i=N+1}^M \frac{|i\rangle\langle i|}{\lambda_i} \\ &= (D^\dagger D)_{oo}^{-1(l)} + (D^\dagger D)_{oo}^{-1(h)} \end{aligned} \quad (5.50)$$

we can also split the source $|D\eta\rangle = D_{oo}^\dagger |\eta'_o\rangle$ and solution into high and low parts,

$$\begin{aligned} |D\eta\rangle &= \sum_{i=0}^N |i\rangle\langle i|D\eta\rangle + \sum_{i=N+1}^M |i\rangle\langle i|D\eta\rangle \\ &= |D\eta\rangle^{(l)} + |D\eta\rangle^{(h)} \end{aligned} \quad (5.51)$$

and similarly for ψ . Using the orthogonality of the eigenvectors the solution is,

$$|\psi\rangle^{(l)} + |\psi\rangle^{(h)} = (D^\dagger D)_{oo}^{-1(l)} |D\eta\rangle^{(l)} + (D^\dagger D)_{oo}^{-1(h)} |D\eta\rangle^{(h)} \quad (5.52)$$

Thus the low mode part of the solution $|\psi\rangle^{(l)}$ can be constructed and we need to solve the equation for $(D^\dagger D)_{oo}^{-1(h)}$ which is better conditioned. If we explicitly remove the low mode part of the source and start with a guess vector orthogonal to the space of low modes (for convenience the initial guess is the source vector $|D\eta\rangle^{(h)}$) the CG algorithm has the property that it will stay orthogonal to the low mode space. Thus we can solve,

$$(D^\dagger D)_{oo}^{(h)} |\psi\rangle^{(h)} = |D\eta\rangle^{(h)} \quad (5.53)$$

without having to project into the high mode space after every multiplication.

5.4.1 Deflation: Results

We have described how deflation works above, now we can ask: when is it beneficial to calculate eigenvectors in order to speed up propagator inversions? That is, if we calculate K eigenvectors, how many solves of $Dx = b$ do we need to do so that the initial investment in calculating K vectors is compensated by the saving from deflation? In the following we use configurations 1900 and 2900 of the $16^3 \times 32 \times 16$, $m = 0.01$, Iwasaki $\beta = 2.13$ with the valence quark mass set to 0.01, very similar results hold for all other configurations. The calculations were carried out on 128 nodes of the IRIDIS High Performance Computing Facility. We gratefully acknowledge the use of the IRIDIS and associated support services at the University of Southampton in the completion of this work. The cost is measured in wall clock time, taking account not only of the number of multiplications but all the other linear algebra that is required for the eigensolver and for deflation. The CG is converged so the residual is less than 10^{-8} and the eigenvectors are calculated with a residual less than 10^{-10} .

In figure 5.10 we show how the cost of convergence of a point source propagator (so twelve CG solves) reduces as we deflate with more and more eigenvectors. In figure 5.11 we show how the cost of calculating eigenvectors increases as we compute more. Deflation undoubtedly speeds up CG convergence, rapidly at first then more slowly as we see less gain from adding more eigenvectors. The question of when the cost of calculating eigenvectors is offset is answered by figure 5.12. This shows the cost (wall clock time) of calculating eigenvectors divided by the time saved from deflating a propagator, so the data point indicates the number of propagators that must be calculated to break even. From the figure this ratio is between one and two for all numbers of eigenvectors up to 100, for example calculating 20 eigenvectors and then doing the deflated CG takes the same amount of time, or less, than the undeflated inversion. Since the eigenvectors can be reused to deflate multiple right hand sides as well as used for low mode averaging we conclude that it is very worthwhile to calculate them in practical simulations, when we often have two or more right hand sides for the same gauge configuration. The deflation technique is very general and will be used in chapter 6, when we have to solve multiple stochastic sources on the same configuration.

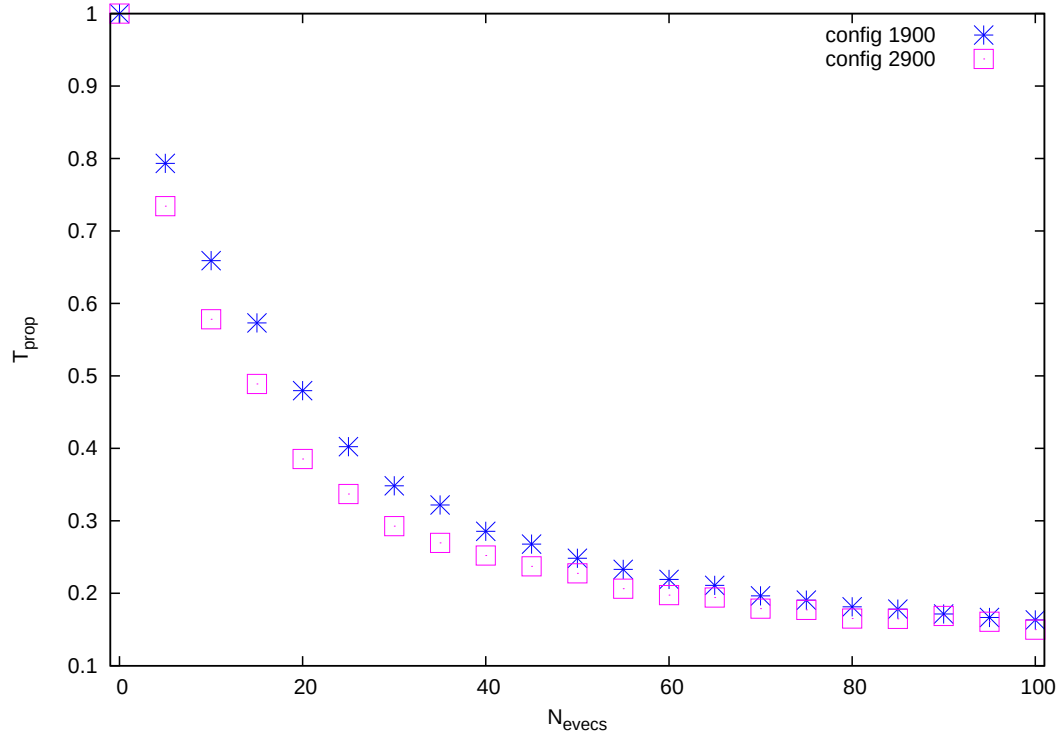


Figure 5.10: The time taken to solve a propagator T_{prop} on the configuration with a point source versus the number of eigenvectors used to deflate the CG. The time taken to calculate the propagator without deflation has been used as a normalization. An approximately $5\times$ speed up is obtained once the eigenvectors have been calculated.

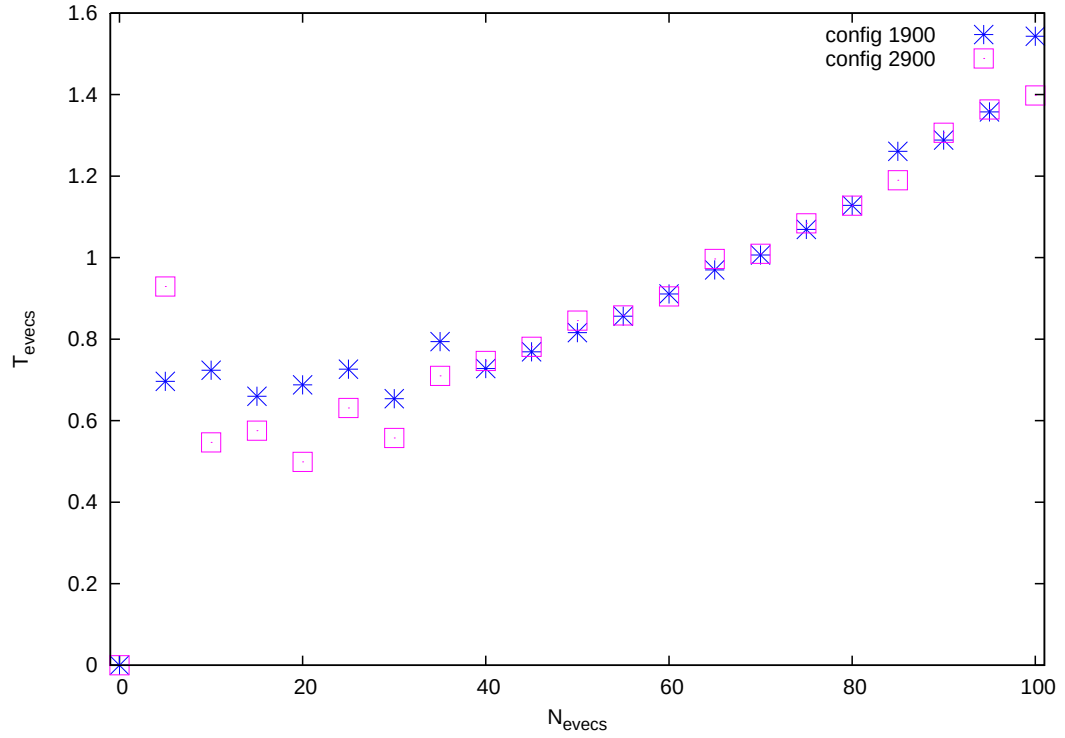


Figure 5.11: The time T_{evecs} taken to calculate N_{evecs} eigenvectors. The time taken to calculate a point source propagator without deflation has been used to normalize the numbers.

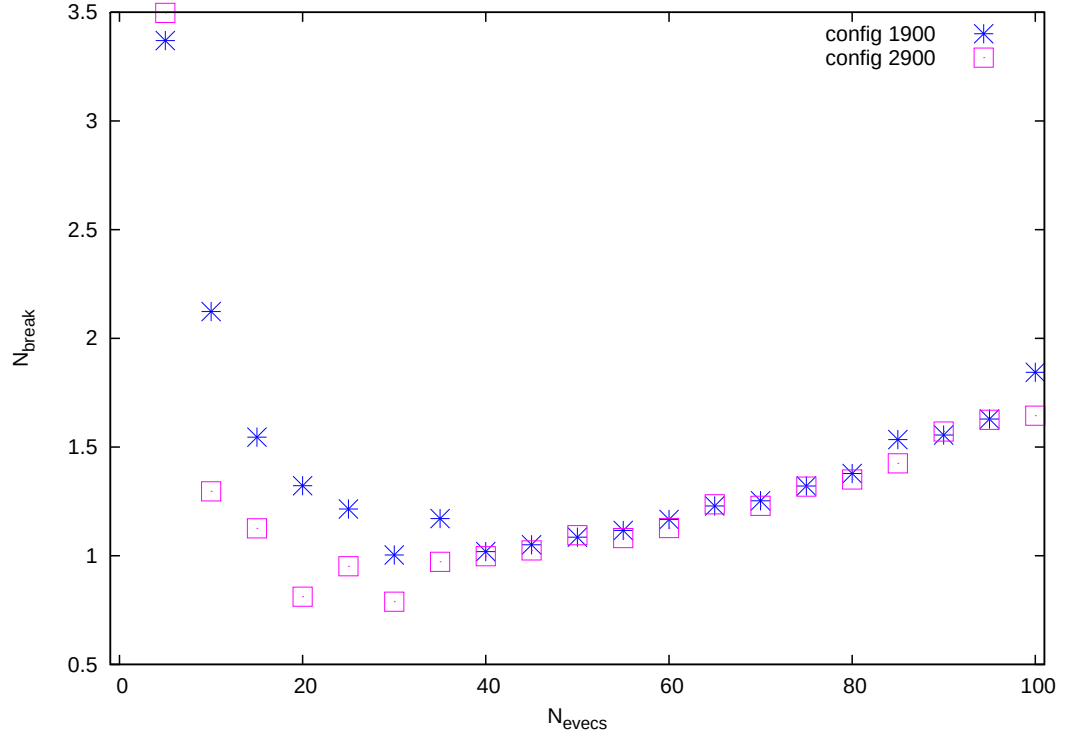
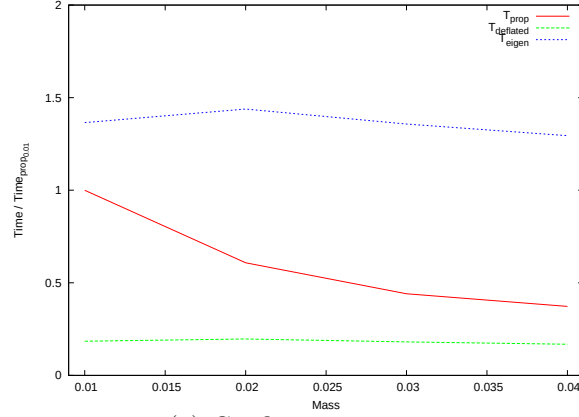
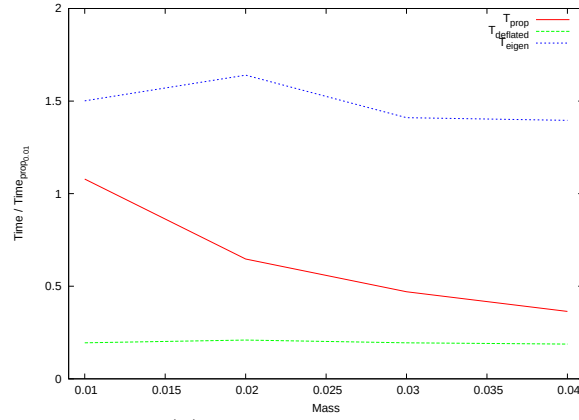


Figure 5.12: The number N_{break} of propagators that must be computed to break even if N_{evecs} eigenvectors have been computed.

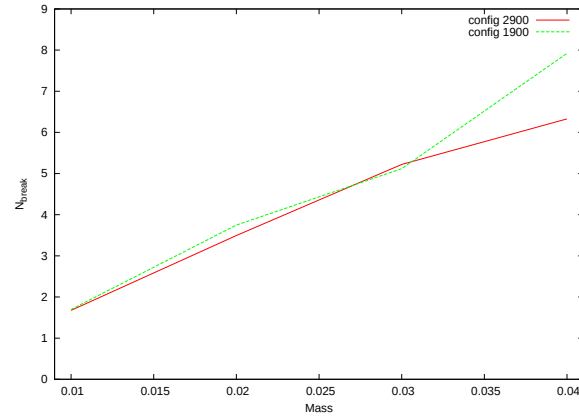
The quark mass affects these results. Figure 5.13(a) and 5.13(b) shows the cost of a point source inversion versus quark mass. The CG cost increases with decreasing mass. The same figure also shows that the cost of calculating eigenvectors, in this case one hundred, is very weakly mass dependant. When these eigenvectors are used to deflate propagator inversions the mass dependence is very much reduced. Figure 5.13(c) shows how many propagators need to be calculated to break even on the cost of one hundred eigenvectors as a function of mass. This increases with increasing mass; the propagator inversion gets cheaper and the effect of low modes is reduced. The lightest mass here corresponds to a pion mass of roughly 430MeV , much greater than the physical pion mass, thus as we reduce the quark masses to get closer to physical hadrons deflation becomes more important and for very light pions the break even number will be very low.



(a) Configuration 2900



(b) Configuration 1900



(c)

Figure 5.13: (a) Time taken to invert a propagator as a function of mass in red, normalized by the time taken to invert the lightest mass propagator. Time taken to calculate 100 eigenvectors in green in the same units and time taken to invert a propagator with deflation. (b) As (a) on configuration 2900. (c) Number of propagators that should be computed to break even on the cost of 100 eigenvectors as a function of mass.

5.5 Instantons

As an interesting test we can examine what happens to the low modes of the Dirac operator in a classical instanton background by integrating the QCD instanton solution. This has been done in [36], [57], [92] and [69], in particular reference [36] examined the domain wall operator in a classical instanton background as a way to investigate how large an L_s was required and a good value for M_5 . In this section we are only concerned to show: that our eigensolver correctly finds near zero modes; that the zero eigenvectors are chiral (see equation 5.63); and that the eigenvectors are strongly spatially correlated with the instanton. This means eigenvectors could be a way to use lattice gauge theory to probe anomaly physics and in particular the η meson.

We follow [57] and [94]. The BPST [18] instanton solution has the gauge field,

$$A_\mu(x) = \frac{x^2}{x^2 + \rho^2} \left(\frac{i}{g} \right) (\partial_\mu S) S^{-1} \quad (5.54)$$

$$S = \frac{x_4 \pm i x_i \sigma_i}{\sqrt{x^2}} \quad (5.55)$$

the plus is for an instanton the minus for an anti-instanton. The path integral, equation 1.15, can be done along a link and gives the regular gauge instanton configuration,

$$U_\mu^{reg} = \exp [i a_\mu(n) \cdot \sigma \phi_\mu(x; \rho)] \quad (5.56)$$

$$\phi_\mu(x; \rho) = \frac{1}{\sqrt{\rho^2 + \sum_{\nu \neq \mu} (x_\nu - c_\nu)^2}} \tan^{-1} \frac{\sqrt{\rho^2 + \sum_{\nu \neq \mu} (x_\nu - c_\nu)^2}}{\rho^2 + \sum_\nu (x_\nu - c_\nu)^2 + (x_\nu - c_\nu)} \quad (5.57)$$

with the instanton radius ρ and centre c_ν . Define the coordinate

$$y_\nu = x_\nu - c_\nu \quad (5.58)$$

then,

$$\begin{aligned}
a_1 &= (-y_4, y_3, -y_2) \\
a_2 &= (-y_3, -y_4, y_1) \\
a_3 &= (y_2, -y_1, -y_4) \\
a_4 &= (y_1, y_2, y_3).
\end{aligned} \tag{5.59}$$

The singular gauge instanton is found via the gauge transformation $g(x) = (x_4 + i\sigma_i x_i)/|x|$,

$$U_\mu^{sing} = \exp [ib_\mu(n) \cdot \sigma(\phi_\mu(x; 0) - \phi_\mu(x; \rho))] \tag{5.60}$$

$$\begin{aligned}
b_1 &= (y_4, y_3, -y_2) \\
b_2 &= (-y_3, y_4, y_1) \\
b_3 &= (y_2, -y_1, y_4) \\
b_4 &= (-y_1, -y_2, -y_3)
\end{aligned} \tag{5.61}$$

doing a parity transform $x_i \rightarrow -x_i$ exchanges instantons and anti-instantons. The singular gauge instanton cannot have its centre on a lattice site but in this case the density is localized around c_ν . We use the singular gauge in order to make the localization of eigenmodes around points of high gauge field density more obvious.

We can then calculate eigenvalues and eigenvectors on this background using the Lanczos algorithm. The low modes of $H_{ov} = \gamma_5 D_{ov}$ are the most physical and easiest to interpret. A zero mode of an operator D that satisfies the Ginsparg-Wilson relation 1.72 obeys the following equation,

$$\begin{aligned}
\gamma_5 D_{ov} v_0 &= 0 \\
(2D\gamma_5 D - D_{ov}\gamma_5)v_0 &= 0 \\
D_{ov}(\gamma_5 v_0) &= 0
\end{aligned} \tag{5.62}$$

implying that, on the subspace of zero modes, γ_5 commutes with the Dirac op-

erator and hence they can be simultaneously diagonalized. Since the eigenvalues of γ_5 are ± 1 the zero modes satisfy,

$$\gamma_5 v_0 = \pm v_0 \quad (5.63)$$

ie. they are chiral. The index theorem [11] relates the number of zero modes to the topological charge via (see [70]),

$$Q_{top} = \frac{1}{32\pi^2} \int d^4x \epsilon_{\mu\nu\rho\sigma} \text{tr}[F_{\mu\nu} F_{\rho\sigma}] = n_- - n_+ \quad (5.64)$$

where $n_{\pm}$ is the number of zero modes with chirality $\pm$. For a single instanton solution the integral gives $Q_{top} = 1$. In the dilute approximation, where the instantons are very far apart, we can just multiply the gauge fields together and every instanton contributes +1 and an anti-instanton contributes -1, which we should find reflected in the eigenvalue spectrum for negative and positive chirality zero modes.

The singular gauge instanton above was implemented. Figure 5.14(a) shows the action density $F_{\mu\nu}(x)F_{\mu\nu}(x)$ of an instanton field calculated using the clover definition [46] of the field strength. It is centred at (3.5, 3.5, 3.5, 3.5) with $\rho = 3.0$. After 10 hits of APE smearing with smearing parameter 0.5 [26], the topological charge is 1. This is measured from the field strength tensor using the methods of [22]. The lowest mode of the operator H_{ov} with $M_5 = 1$, $m = 0.0$ and $L_s = 16$ was found and its density $v^\dagger v(x)$ is plotted in figure 5.14(b). The lowest eigenvalue is $\lambda_0 = -7 \times 10^{-8}$ with the next highest $\lambda_1 = 0.3173$ and the chirality of the mode is $\chi = 1.0$. We have the correct result according to the index theorem and importantly, as figure 5.14 shows, the low mode is spatially congruent with the gauge field. Figure 5.14(c) shows this spatial correlation holds even at non-zero mass where the lowest mode is found to be $\lambda_0^{m=0.01} = 0.01$. The higher modes show much less correlation with the instanton, see figure 5.14(d). From the plots in figure 5.15 we see this also holds for more than one instanton.

The low modes are strongly correlated in space and time with the instantons, the higher modes less so. For correlation functions dominated by the interaction of the quark field with the gauge background we expect that the low modes contain most of the relevant information and including only the lowest few should be enough. How many “a few” is, how this depends on the correlation function and

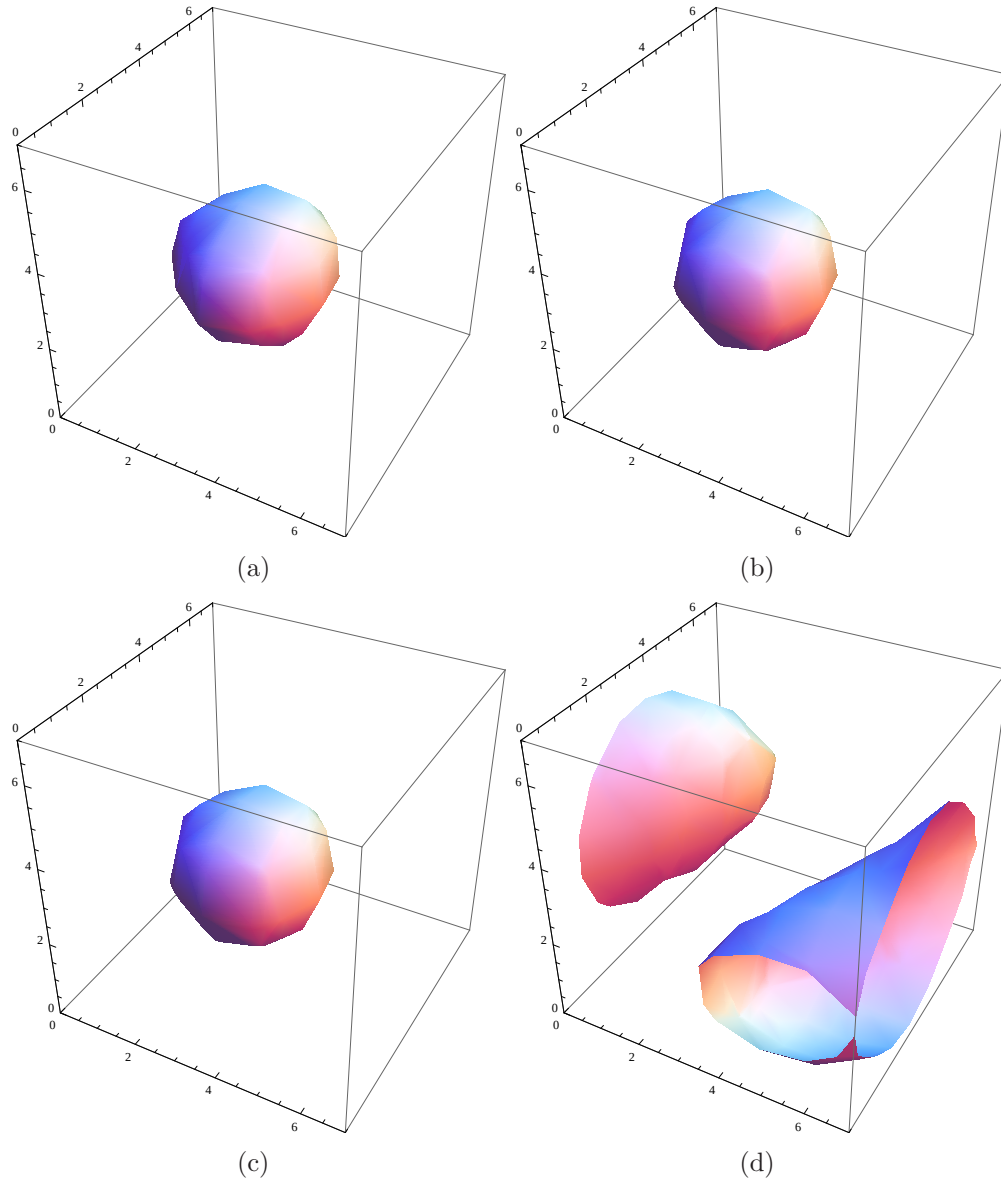
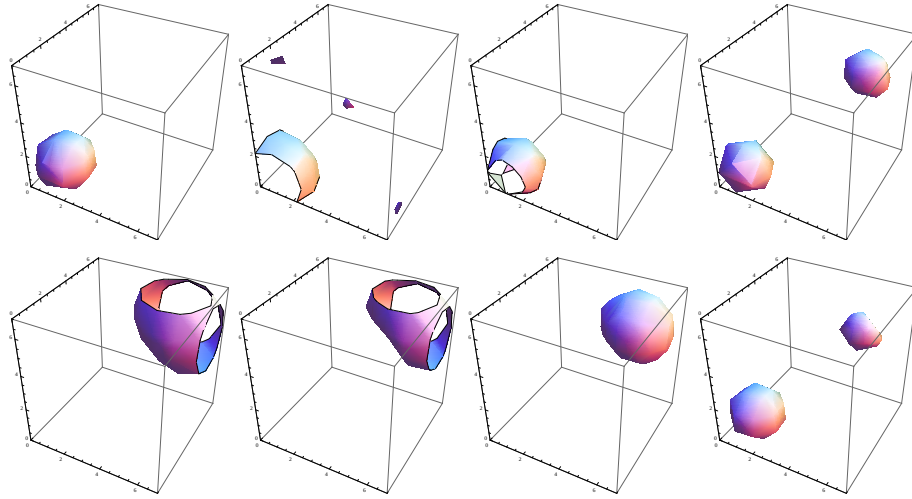
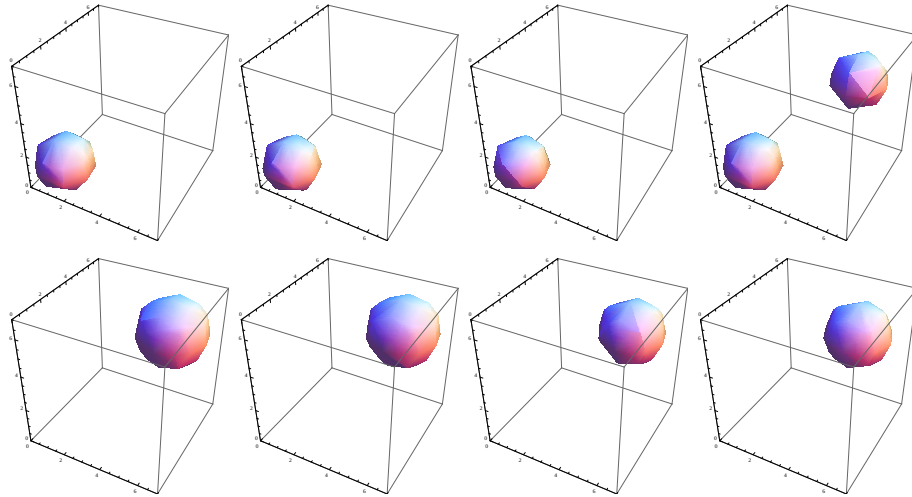


Figure 5.14: Single instanton on a 8^4 lattice, all the plots show a contour plot of the density in xyz at timestep 5. (a) Action density. (b) Density $v^\dagger v(x)$ of the lowest eigenmode of the massless overlap operator H_{ov} with $L_s = 16$. (c) Lowest mode of the massive, $m = 0.01$, overlap operator, $L_s = 16$. (d) Third lowest mode of the massless overlap.



Action density



Low mode density

Figure 5.15: Action density for a two instanton configuration and the density of the first eigenvalue of the massless overlap operator with $L_s = 16$. The plots show a constant density contour in xyz for each of the eight timesteps and the instantons are centered at $(1.5, 1.5, 1.5, 1.5)$ and $(5.5, 5.5, 5.5, 5.5)$.

how we may account for the high modes we left out is dealt with in the next chapter.

5.6 Conclusions

We have developed a code that can find the eigenvalues and eigenvectors of Hermitian Dirac operators. Chebyshev acceleration makes the computation of eigenvectors of $D_{oo}^\dagger D_{oo}$ quite affordable, with the right parameters it is $O(5)$ times faster than without Chebyshev acceleration. When using Chebyshev filters we are trading more multiplications for less linear algebra. On the next generation of supercomputers (eg. the IBM Bluegene/Q) multiplication by the Dirac operator can be performed extremely efficiently, essentially at the cost of loading a vector from memory. Thus spending more time in the multiplication phase of the algorithm to reduce time spent elsewhere is a good strategy for the future. The eigenvectors can be used for deflation of the conjugate gradient solver. Since solving the system $Dx = b$ is a large part of the budget of most lattice simulations a means to accelerate this is very desirable. Low mode deflation accomplishes just this. Further we have shown the cost of calculating a given number of low modes is virtually independent of quark mass while the cost of a conjugate gradient solve increases rapidly with decreasing mass. Thus as quark masses get smaller deflation becomes crucial. We expect that deflation for solving multiple right hand sides will be used in future calculations of the RCB-UKQCD collaboration.

Chapter 6

Low Mode Averaging

The renormalization group introduced the idea that low and high energy physics can decouple. In lattice simulations we have cut-offs of the order of 2GeV but it is the low energy dynamics that are responsible for low energy observables. The propagator, calculated from the inverse Dirac matrix, contains contributions from all modes up to the cut-off. It might then be reasonable to do calculations very accurately in the low mode sector and less accurately in the high mode sector and expect that, since the low mode sector is most important, high accuracy there translates to high accuracy in the final result. This is dependent on the observable in question. Below we will describe several methods of low mode averaging and choose an optimal one as well as explore in what situations we have low mode dominance.

6.1 Low Mode Averaging with Domain Wall Fermions

We will be concerned with evaluating connected correlation functions

$$C(t) = \sum_{\vec{x}, \vec{y}, \tau} \text{Tr} \left[S^{(q)}(x, y) \Gamma_1 S^{(q')}(y, x) \Gamma_2 \right] \quad (6.1)$$

and disconnected loops,

$$D(t) = \sum_{\vec{x}} \text{Tr} \left[S^{(q)}(x, x) \Gamma \right]. \quad (6.2)$$

In these equations $x = (\vec{x}, t)$ and $y = (\vec{y}, t + \tau)$. The technique of low mode averaging is not restricted to these examples at all but this is the simplest situation for a first study.

The volume average in 6.1 and 6.2 cannot be done using point sources, there are simply too many points on the lattice. Low mode averaging is the technique of doing this volume average exactly in a small dimensional subspace [47], [72]. The subspace is constructed from the lowest modes of some operator and these low modes are supposed to overlap strongly with those of $S(x, y)$. Several choices of operator are possible which are explored below. In general we will find,

$$S(x, y) = S_L(x, y) + S_H(x, y) \quad (6.3)$$

This gives,

$$C = C_{LL} + C_{HL} + C_{LH} + C_{HH} \quad (6.4)$$

$$D = D_L + D_H \quad (6.5)$$

where,

$$C_{HL}(t) = \sum_{\vec{x}, \vec{y}, \tau} \text{Tr} \left[S_H^{(q)}(x, y) \Gamma_1 S_L^{(q')}(y, x) \Gamma_2 \right] \quad (6.6)$$

and similarly for the other terms. We will find that the low mode propagator can be written as the outer product of two vectors,

$$S_L^{(q)}(x, y) = \sum_i^N |a_i^{(q)}(x)\rangle \langle b_i^{(q)}(y)| \quad (6.7)$$

where a_i and b_i will be simply related to the i^{th} eigenvector of some operator and as the number of eigenvectors, N , tends to the total number of eigenvectors of the operator, M , the lowmode propagator tends to the full propagator, $S_L \rightarrow S$. In the following subsections we will give a_i and b_i for various operators. C_{LL} and

D_L are constructed as follows,

$$\begin{aligned} C_{LL} &= \sum_{\vec{x}, \vec{y}, \tau} \sum_{i,j}^N \text{Tr} \left[|a_i^{(q)}(x)\rangle \langle b_i^{(q)}(y)| \Gamma_1 |a_j^{(q')}(y)\rangle \langle b_j^{(q')}(x)| \Gamma_2 \right] \\ &= \sum_{\vec{x}, \vec{y}, \tau} \sum_{i,j}^N \text{Tr} \left[\langle b_j^{(q')}(x)| \Gamma_2 |a_i^{(q)}(x)\rangle \langle b_i^{(q)}(y)| \Gamma_1 |a_j^{(q')}(y)\rangle \right] \end{aligned} \quad (6.8)$$

$$\begin{aligned} D_L &= \sum_{\vec{x}} \sum_i^N \text{Tr} \left[|a_i^{(q)}(x)\rangle \langle b_i^{(q)}(x)| \Gamma \right] \\ &= \sum_{\vec{x}} \sum_i^N \text{Tr} \left[\langle b_i^{(q)}(x)| \Gamma |a_i^{(q)}(x)\rangle \right]. \end{aligned} \quad (6.9)$$

We will discuss the high mode pieces in the next section.

6.1.1 Effective Overlap Operator

From chapter 1 equation 1.102 we recall that we can write the quark propagator in terms of an effective overlap operator,

$$D_{ov} = [\mathcal{P}^{-1} D_{DWF}^{-1}(m=1) D_{DWF}(m) \mathcal{P}]_{11}$$

$$S = \tilde{D}_{ov}^{-1} = \frac{1}{1-m} [D_{ov}^{-1} - 1].$$

Since the Shamir kernel H_T , equation 1.86, is Hermitian so is the transfer matrix T , equation 1.85, and hence,

$$H_{ov} = \gamma_5 D_{ov} \quad (6.10)$$

is Hermitian and therefore its eigenvalues are real and its eigenvectors are mutually orthogonal.

We can calculate eigenvectors of H_{ov} by the methods of chapter 5 and the inverse can be constructed from them in the standard way,

$$H_{ov} |i\rangle = \gamma_5 D_{ov} |i\rangle = \lambda_i |i\rangle \quad (6.11)$$

$$D_{ov}^{-1} = \sum_i^M \frac{|i\rangle \langle i| \gamma_5}{\lambda_i} \quad (6.12)$$

$$S(x, y) = \frac{1}{1-m} \sum_i^M |i(x)\rangle \langle i(y)| \left(\frac{\gamma_5}{\lambda_i} - 1 \right). \quad (6.13)$$

Truncating this sum at $N < M$ gives the low mode contribution to the propagator $S_L(x, y)$ and we can write, following equation 6.7,

$$\begin{aligned} |a_i(x)\rangle &= \frac{1}{1-m} |i(x)\rangle \\ \langle b_i(x)| &= \langle i(y)| \left(\frac{\gamma_5}{\lambda_i} - 1 \right). \end{aligned} \quad (6.14)$$

6.1.2 Hermitian DWF Operator

The operator,

$$H_{DWF} = \gamma_5 \mathcal{R}_5 D_{DWF} \quad (6.15)$$

is Hermitian. Let,

$$H_{DWF} |i(s)\rangle = \gamma_5 \mathcal{R}_5 D_{DWF} |i(s)\rangle = \lambda_i |i(s)\rangle \quad (6.16)$$

where we have explicitly written out the fifth spatial index, s , of the eigenvector which is now five dimensional.

$$D_{DWF}^{-1} = \sum_i^M \frac{|i(s)\rangle \langle i(L_s - 1 - t)| \gamma_5}{\lambda_i} \quad (6.17)$$

and the equation, 1.48,

$$\begin{aligned} S &= P_R D_{DWF}^{-1}(L_s - 1, 0) P_R + P_R D_{DWF}^{-1}(L_s - 1, L_s - 1) P_L \\ &\quad + P_L D_{DWF}^{-1}(0, 0) P_R + P_L D_{DWF}^{-1}(0, L_s - 1) P_L, \end{aligned}$$

implies,

$$\begin{aligned} S &= \sum_i \frac{1}{\lambda_i} (P_R |i(L_s - 1)\rangle + P_L |i(0)\rangle) \\ &\quad \times (\langle i(L_s - 1)| P_R - \langle i(0)| P_L) \end{aligned} \quad (6.18)$$

so the a and b required for equation 6.7 are,

$$\begin{aligned} |a_i(x)\rangle &= P_R|i(x, L_s - 1)\rangle + P_L|i(x, 0)\rangle \\ \langle b_i(x)| &= \frac{1}{\lambda_i} (-\langle i(x, L_s - 1)|P_R + \langle i(x, 0)|P_L). \end{aligned} \quad (6.19)$$

Note the important point that we only need the first and last components of the eigenvectors for LMA. This makes the storage of five dimensional eigenvectors of H_{DWF} only a factor two more expensive than for $4d$, not a factor L_s .

6.1.3 Preconditioned DWF

We can also use eigenvectors of the preconditioned Dirac operator D_{oo} . Referring to the Schur decomposition, equation 3.32,

$$\begin{aligned} D_{DWF} &= \begin{pmatrix} M_{ee} & M_{eo} \\ M_{oe} & M_{oo} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ M_{oe}M_{ee}^{-1} & 1 \end{pmatrix} \begin{pmatrix} M_{ee} & 0 \\ 0 & D_{oo} \end{pmatrix} \begin{pmatrix} 1 & M_{ee}^{-1}M_{eo} \\ 0 & 1 \end{pmatrix} \\ &= LDU \\ D_{oo} &= M_{oo} - M_{oe}M_{ee}^{-1}M_{eo}. \end{aligned}$$

We will be using four dimensional preconditioning, this makes the following analysis slightly simpler and also for four dimensional preconditioning we can use the Chebyshev shifting trick of section 5.2.4. The Hermitian operator $D_{oo}^\dagger D_{oo}$ is used,

$$D_{oo}^\dagger D_{oo}|i\rangle = \lambda_i|i\rangle \quad (6.20)$$

$$D_{oo}^{-1} = \sum_i^N \frac{|i\rangle\langle i|D_{oo}^\dagger}{\lambda_i}. \quad (6.21)$$

From section 3.2.1 the solution to $D_{DWF}\psi = \eta$ can be written in terms of even and odd subspaces,

$$\begin{aligned} |\psi_o\rangle &= D_{oo}^{-1}|\eta'_o\rangle \\ &= D_{oo}^{-1}|\eta_o\rangle - D_{oo}^{-1}M_{oe}M_{ee}^{-1}|\eta_e\rangle \\ |\psi_e\rangle &= M_{ee}^{-1}|\eta_e\rangle - M_{ee}^{-1}M_{eo}|\psi_o\rangle \\ &= (M_{ee}^{-1} + M_{ee}^{-1}M_{eo}D_{oo}^{-1}M_{oe}M_{ee}^{-1})|\eta_e\rangle - M_{ee}^{-1}M_{eo}D_{oo}^{-1}|\eta_o\rangle. \end{aligned} \quad (6.22)$$

Then inserting the expansion of D_{oo}^{-1} gives,

$$\begin{pmatrix} |\psi_e\rangle \\ |\psi_o\rangle \end{pmatrix} = \begin{pmatrix} S_{ee} & S_{eo} \\ S_{oe} & S_{oo} \end{pmatrix} \begin{pmatrix} |\eta_e\rangle \\ |\eta_o\rangle \end{pmatrix} \quad (6.23)$$

with

$$\begin{aligned} S_{ee} &= M_{ee}^{-1} + \sum_i M_{ee}^{-1} M_{eo} \frac{|i\rangle\langle i|}{\lambda_i} D_{oo}^\dagger M_{oe} M_{ee}^{-1} \\ S_{eo} &= \sum_i -M_{ee}^{-1} M_{eo} \frac{|i\rangle\langle i|}{\lambda_i} D_{oo}^\dagger \\ S_{oe} &= \sum_i -\frac{|i\rangle\langle i|}{\lambda_i} D_{oo}^\dagger M_{oe} M_{ee}^{-1} \\ S_{oo} &= \sum_i \frac{|i\rangle\langle i|}{\lambda_i} D_{oo}^\dagger \end{aligned} \quad (6.24)$$

Note how indices $\{e, o\}$ always match in these expressions.

Low mode averaging with eigenvectors of $D_{oo}^\dagger D_{oo}$ is a little more complicated than previously because D_{DWF}^{-1} is a 2×2 checkerboarded matrix. As in the previous subsection the propagator is,

$$\begin{aligned} S &= P_R D_{DWF}^{-1}(L_s - 1, 0) P_R + P_R D_{DWF}^{-1}(L_s - 1, L_s - 1) P_L \\ &\quad + P_L D_{DWF}^{-1}(0, 0) P_R + P_L D_{DWF}^{-1}(0, L_s - 1) P_L. \end{aligned}$$

If say, $x = x_e$ is even and $y = y_o$ is odd then $D^{-1}(x_e, s; y_o, t) = S_{eo}(x_e, s; y_o, t)$, since we have used four dimensional preconditioning the value of the fifth index doesn't affect the parity. Writing,

$$\begin{aligned} S_{eo}(x_e, s; y_o, t) &= \sum_i \frac{1}{\lambda_i} (-M_{ee}^{-1} M_{eo} |i(x_e)\rangle)_s (\langle i(y_o) | D_{oo}^\dagger)_t \\ &= \sum_i |c_i(x_e, s)\rangle \langle d_i(y_o, t)| \end{aligned} \quad (6.25)$$

gives,

$$S(x_e, y_o) = (P_R |c_i(x_e, L_s - 1)\rangle + P_L |c_i(x_e, 0)\rangle) \times (\langle d_i(y_o, 0) | P_R + \langle d_i(y_o, L_s - 1) | P_L). \quad (6.26)$$

So,

$$\begin{aligned} |a_i^{eo}(x)\rangle &= P_R |c_i(x_e, L_s - 1)\rangle + P_L |c_i(x_e, 0)\rangle \\ \langle b_i^{eo}(x)| &= \langle d_i(y_o, 0) | P_R + \langle d_i(y_o, L_s - 1) | P_L \end{aligned} \quad (6.27)$$

gives $S_L(x_e, y_o)$ as in equation 6.7. The terms S_{oe} and S_{oo} contribute in a similar

way. S_{ee} has two terms to deal with. We know M_{ee}^{-1} and have given an explicit expression for it in section 3.2.1, which we use when forming correlation functions.

The correlation function becomes,

$$C(t) = \sum_{\vec{x}, \vec{y}, \tau} C(x, y)$$

$$C(x, y) = \begin{cases} \text{Tr}[S_{ee}^{(q)}(x, y)\Gamma_1 S_{ee}^{(q')}(y, x)\Gamma_2] & : \quad x \text{ even}, y \text{ even} \\ \text{Tr}[S_{eo}^{(q)}(x, y)\Gamma_1 S_{oe}^{(q')}(y, x)\Gamma_2] & : \quad x \text{ even}, y \text{ odd} \\ \text{Tr}[S_{oe}^{(q)}(x, y)\Gamma_1 S_{eo}^{(q')}(y, x)\Gamma_2] & : \quad x \text{ odd}, y \text{ even} \\ \text{Tr}[S_{oo}^{(q)}(x, y)\Gamma_1 S_{oo}^{(q')}(y, x)\Gamma_2] & : \quad x \text{ odd}, y \text{ odd} \end{cases}$$

and the disconnected piece,

$$D(t) = \sum_{\vec{x}} D(x) \tag{6.28}$$

$$D(x) = \begin{cases} \text{Tr}[S_{ee}^{(q)}(x, x)\Gamma] & : \quad x \text{ even} \\ \text{Tr}[S_{oo}^{(q)}(x, x)\Gamma] & : \quad x \text{ odd} \end{cases}$$

6.2 Estimating the High Mode Part

If low mode averaging is to be effective the low mode parts, C_{LL} and D_L , should be dominant. However we find that at short time separations the high mode pieces still contribute and it is necessary to estimate them. This should be done fairly cheaply, using a small number of inversions, which can be sped up by deflating with the already calculated eigenvectors, see section 5.4. We consider point sources and complex $\mathbb{Z}(2)$ noise sources.

6.2.1 Point Sources

The connected correlation function can be estimated with point source propagators,

$$C(t) \approx C^{point}(t) = \sum_{\vec{x}} \text{Tr} \left[S^{(q)}(x, 0)\Gamma_1 S^{(q')}(0, x)\Gamma_2 \right]. \tag{6.29}$$

The disconnected piece can also be estimated with point source propagators but in this case the signal is extremely noisy and the method is not useful. We give a better way to calculate $D(t)$ in the next section. C^{point} will have a larger variance than C_{LL} since we aren't including as many samples and relying on translational invariance. However it does include the contribution from the high modes since $S(x, 0)$ is obtained from solving the full Dirac equation.

C^{point} has a mode decomposition too,

$$S(x, 0) = S_L(x, 0) + S_H(x, 0) \quad (6.30)$$

and gives,

$$C^{point} = C_{LL}^{point} + C_{HL}^{point} + C_{LH}^{point} + C_{HH}^{point} \quad (6.31)$$

with

$$C_{HL}^{point}(t) = \sum_{\vec{x}} Tr \left[S_H^{(q)}(x, 0) \Gamma_1 S_L^{(q')}(0, x) \Gamma_2 \right], \quad (6.32)$$

and similarly for the other terms. We have shown how to calculate C_{LL} above for various different operators, C_{LL}^{point} can be constructed the same way by substituting,

$$S_L^{(q)}(x, 0) = \sum_i^N |a_i^{(q)}(x)\rangle \langle b_i^{(q)}(0)|. \quad (6.33)$$

Then simply notice, [13],

$$C^{point} - C_{LL}^{point} = C_{HL}^{point} + C_{LH}^{point} + C_{HH}^{point}. \quad (6.34)$$

So if we calculate the correlation function with point sources, reusing existing code and subtract C_{LL}^{point} , which we construct in tandem with C_{LL} , we have an estimate for the high mode piece,

$$C_{HL} + C_{LH} + C_{HH} \approx C_{HL}^{point} + C_{LH}^{point} + C_{HH}^{point}. \quad (6.35)$$

Thus,

$$C(t) \approx C^{LMA}(t) = C_{LL}(t) + C^{point}(t) - C_{LL}^{point}(t). \quad (6.36)$$

6.3 Stochastic Sources

For disconnected diagrams we need to compute terms like $\sum_{\vec{x}} \text{Tr}[S(x, x)\Gamma]$ which require a sum over volume. Using point sources here, although equivalent in the limit of many gauge configurations, gives a very noisy estimate of this quantity. A way to do these volume sums using the “all-to-all” propagator $S(x, y)$ is via stochastic sources [49], [131], [64]. K sources $\eta_k(x)$ are constructed with the property,

$$\frac{1}{K} \sum_k^K |\eta_k(x)\rangle \langle \eta_k(y)| \rightarrow \delta(x, y) \quad (6.37)$$

for large K . Reference [49] suggests that $\mathbb{Z}(2) \otimes \mathbb{Z}(2)$ random numbers satisfy equation 6.37 optimally for a given K , so we use these. That is we randomly assign to each complex variable in η_k one of the four values,

$$\pm \frac{1}{\sqrt{2}} \pm \frac{i}{\sqrt{2}}. \quad (6.38)$$

We then solve the Dirac equation for this source vector to get $\psi_k(x)$ and

$$\frac{1}{K} \sum_k^K |\psi_k(x)\rangle \langle \eta_k(y)| = \frac{1}{K} \sum_k^K S(x, z) |\eta_k(z)\rangle \langle \eta_k(y)| \rightarrow S(x, y). \quad (6.39)$$

The error in this construction is caused by the deviation of equation 6.37 from a delta function at finite K .

6.3.1 Connected Correlation Functions

For connected correlation functions we use the “one end trick” [104]. Take

$$C(t) = \frac{1}{V} \sum_{\vec{x}, \vec{y}} \text{Tr} \left[S^{(q')}(x, y) \Gamma_1 S^{(q)}(y, x) \Gamma_2 \right] \quad (6.40)$$

which is summed over space but not time, τ . Insert a delta function, use $S(x, y) = \gamma_5 S^\dagger(y, x) \gamma_5$ and rearrange the trace to,

$$C(t) = \frac{1}{V} \sum_{\vec{x}, \vec{y}, \vec{z}} \text{Tr} \left[\gamma_5 \Gamma_2 S^{(q')}(y, x) \delta(\vec{x}, \vec{z}) \Gamma_1 \gamma_5 S^{(q)}(y, z) \right]. \quad (6.41)$$

We construct stochastic sources that are nonzero only on a timeslice τ ,

$$\eta(\vec{x}, \tau) = \begin{cases} \mathbb{Z}(2) \otimes \mathbb{Z}(2) & \text{if } t_x = \tau \\ 0 & \text{otherwise} \end{cases} \quad (6.42)$$

then

$$\frac{1}{K} \sum_k^K |\eta_k(\vec{x}, \tau)\rangle \langle \eta_k(\vec{y}, \tau)| \rightarrow \delta(\vec{x}, \vec{y}) \delta(\tau, t_y) \quad (6.43)$$

so we have explicitly put in the delta function in time. Inserting the stochastic approximation to the delta function we obtain,

$$C(t) = \frac{1}{VK} \sum_{\vec{x}, \vec{y}, \vec{z}, k} Tr \left[\gamma_5 \Gamma_2 S^{(q')}(y, x) |\eta_k(x)\rangle \langle \eta_k(z)| \Gamma_1 \gamma_5 S^{\dagger(q)}(y, z) \right]. \quad (6.44)$$

Thus,

$$C(t) = \frac{1}{VK} \sum_{\vec{y}, k} Tr [\langle \psi_\Gamma(y) | \gamma_5 \Gamma_2 | \psi(y) \rangle] \quad (6.45)$$

$$|\psi_k(y)\rangle = S^{(q')}(y, x) |\eta_k(x)\rangle \quad (6.46)$$

$$|\psi_{k,\Gamma}(y)\rangle = S^{(q')}(y, x) (\Gamma_1 \gamma_5)^\dagger |\eta_k(x)\rangle. \quad (6.47)$$

This trick greatly reduces the noise by inserting only one noisy delta function compared to the naive replacment $S \rightarrow D^{-1} |\eta\rangle \langle \eta|$ for each propagator, which uses two. Note that $\Gamma_1 = \gamma_5 \rightarrow \psi = \psi_\Gamma$, thus with one inversion of the Dirac equation we can estimate the correlation function for the pseudoscalar operator. This does however allow us to make a good estimate of the pion mass quite cheaply and we will use this method later, when accuracy in the connected correlation function is not crucial and only the pion is needed. Reference [32] calls this a *Z2PSWall* source.

For each of the sixteen independent gamma matrices Γ , ψ_Γ is different. We can avoid doing sixteen inversions by computing the spin structure explicitly. We take each spin component of our noise vectors, η_α ,

$$\eta(x) = \sum_{\alpha=0}^3 \eta_\alpha(x) \quad (6.48)$$

and solve each η_α separately to make ψ_α , before summing all four to get the

final answer. Reference [63] calls this process of splitting noise vectors into smaller subsets and solving for each element “dilution”. The sources above are diluted in spin. The one end trick sources are diluted in time, we may repeat the measurement above for $\tau = \{0, \dots, T - 1\}$. We can also dilute in colour or any spatial direction. In the limit of diluting all variables we are evaluating the exact all-to-all propagator. We choose the variable to dilute by trying to avoid relying on $O(1)$ cancellations over many noise samples. The propagator falls exponentially with separation so, for large separations, we are trying to find an exponentially small signal amongst $O(1)$ random noise. In this case dilution in time is essential.

6.3.2 Disconnected Correlation Functions

For disconnected correlation functions we cannot use the one end trick and have to rely on the substitution

$$S(x, y) \approx D^{-1}(x, z) |\eta(z)\rangle \langle \eta(y)| = |\psi(x)\rangle \langle \eta(y)|. \quad (6.49)$$

The disconnected correlator is

$$D(t) = \frac{1}{VT} \sum_{\vec{x}, \vec{y}, \tau} \text{Tr} [S(x, x) \Gamma] \text{Tr} [S(y, y) \Gamma]$$

the two traces can be evaluated separately. Inserting equation 6.50 they become,

$$\text{Tr} [S(x, x) \Gamma] \approx \frac{1}{K} \sum_k \text{Tr} [\langle \eta_k(x) | \Gamma | \psi_k(x) \rangle]. \quad (6.50)$$

In this thesis we will be most concerned with $\Gamma = \gamma_5$. In this case the sum over the whole lattice volume,

$$\sum_x \text{Tr} [S(x, x) \gamma_5] \quad (6.51)$$

gives the topological index, if we have a Dirac operator obeying the Ginsparg-Wilson relation, [70], and vanishes after a gauge average. This means for the sum over the spatial volume, we are looking for a very small signal to cancel amongst $O(1)$ noise. It is vital to dilute in spin and not rely on the stochastic average

here. Any other dilutions are of lesser importance.

6.3.3 LMA with Stochastic Sources

For LMA with stochastic sources we use the same idea: estimate the correlator in the usual way; subtract the low mode piece; and combine this with the low mode average, C_{LL} . The subtraction of the low mode piece from the the stochastic estimate is slightly more difficult.

$$C(t) \approx C^{stoch}(t) = \frac{1}{K} \sum_k \sum_{\vec{x}, \vec{y}, \tau} Tr \left[|\psi_k^{(q)}(x)\rangle \langle \eta_k^{(q)}(y)| \Gamma_1 |\psi_k^{(q')}(y)\rangle \langle \eta_k^{(q')}(x)| \Gamma_2 \right] \quad (6.52)$$

$$D(t) \approx D^{stoch}(t) = \frac{1}{K} \sum_k \sum_{\vec{x}, \vec{y}, \tau} Tr \left[|\psi_k^{(q)}(x)\rangle \langle \eta_k^{(q)}(y)| \Gamma \right] \quad (6.53)$$

where

$$\tilde{D}_{ov} |\psi_k\rangle = |\eta_k\rangle$$

and

$$\frac{1}{K} \sum_k |\eta_k(x)\rangle \langle \eta_k(y)| \rightarrow \delta(x, y)$$

as K becomes large. Then, as in the point source case,

$$\begin{aligned} C(t) &\approx C^{LMA}(t) = C_{LL}(t) + C^{stoch}(t) - C_{LL}^{stoch}(t) \\ D(t) &\approx D^{LMA}(t) = C_L(t) + D^{stoch}(t) - D_L^{stoch}(t). \end{aligned} \quad (6.54)$$

S_L is constructed using equation 6.7,

$$S_L^{stoch}(x, y) = \frac{1}{K} \sum_k \sum_i^N |a_i^{(q)}(x)\rangle \langle b_i^{(q)}(z)| \eta_k(z)\rangle \langle \eta_k(y)|. \quad (6.55)$$

where we take the inverse in the low mode space and multiply with the stochastic approximation to the delta function.

6.4 Meson Correlation functions

In this section we explore low mode averaging of meson correlation functions using each of the three different types of eigenvectors.

6.4.1 Comparison of Eigenvectors

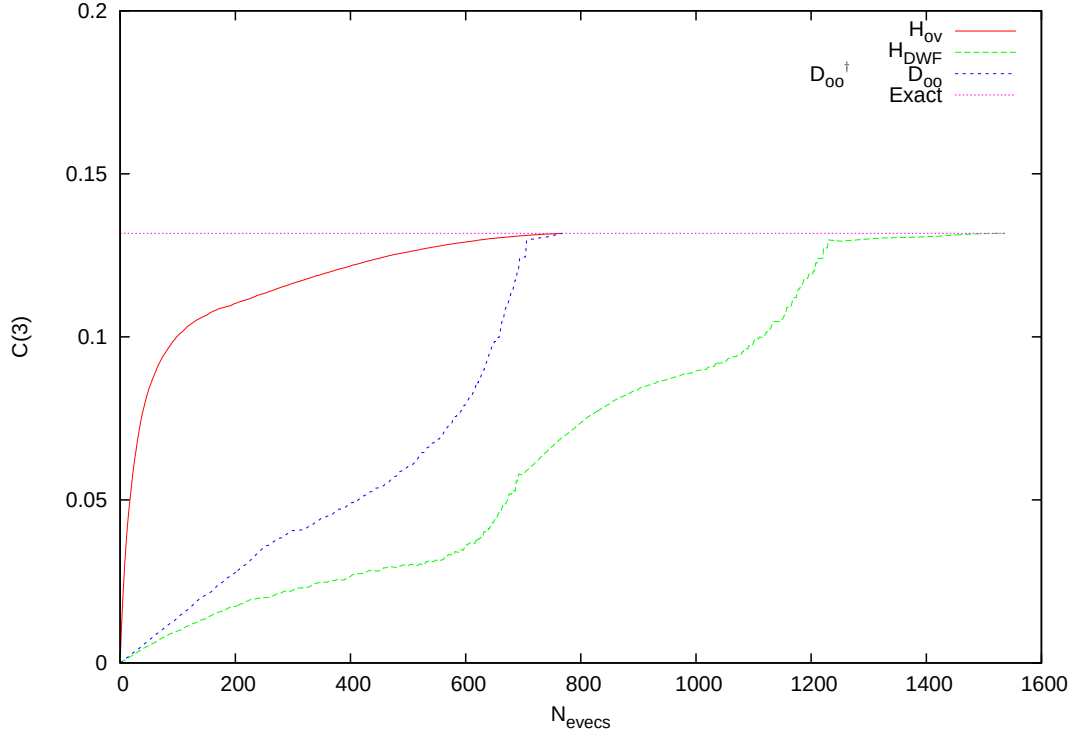


Figure 6.1: Plot showing $C^{LMA}(t=3)$ on a $2 \times 2 \times 4 \times 4$, $L_s = 2$ lattice. The exact value has been calculated by placing a point source on every lattice point and this is compared to the result calculated by summing over N_{evecs} eigenvectors. All three methods reproduce the exact result when all eigenvectors are included.

In order of expense of eigenvectors, $H_{ov} > H_{DWF} > D_{oo}^\dagger D_{oo} \sim 4.7 > 2.3 > 1$. Unless there is a significant advantage of one type of eigenvector over another the reduced cost of $D_{oo}^\dagger D_{oo}$ eigenvectors, as well as their additional usefulness for deflation, makes them preferable. However eigenvectors of H_{ov} and H_{DWF} could still be of use. H_{DWF} eigenvectors cost roughly twice as much to calculate, but for low mode averaging only the first and last components are needed instead of all L_s , thus it becomes easier and cheaper to store large numbers of them. H_{ov}

eigenvectors are four dimensional objects and thus another factor two less storage is necessary, they are also the most physical and may be expected to be better for low mode averaging. A check of a very small test case, $2 \times 2 \times 4$, $L_s = 2$, random gauge, quark mass 0.01, in figure 6.1 indicates that the eigenvalues of H_{ov} smoothly and rapidly saturate the correlation function, eigenvalues of the $5d$ operators show a slower approach to the exact result. However this is far from a realistic situation, only an empirical test can demonstrate which of the three types is preferable.

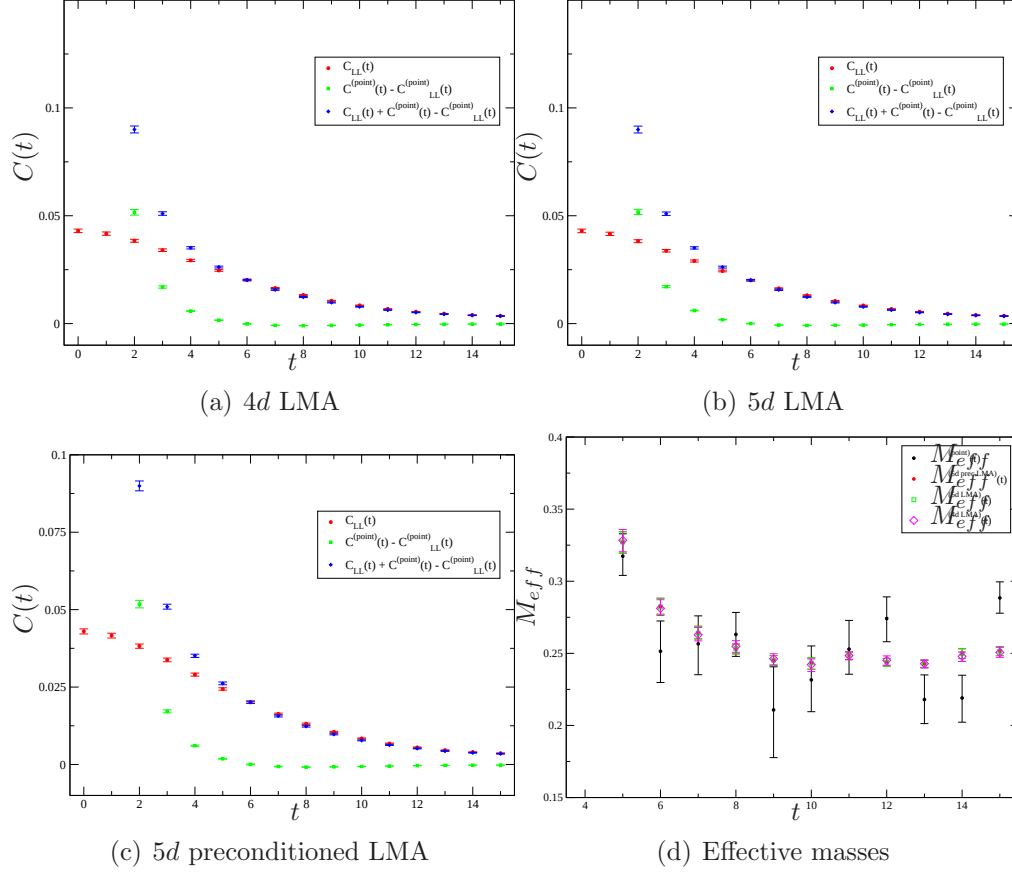
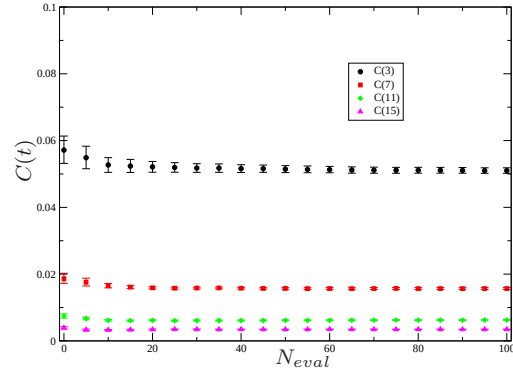
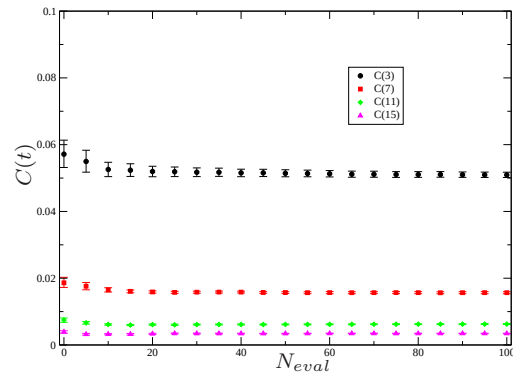


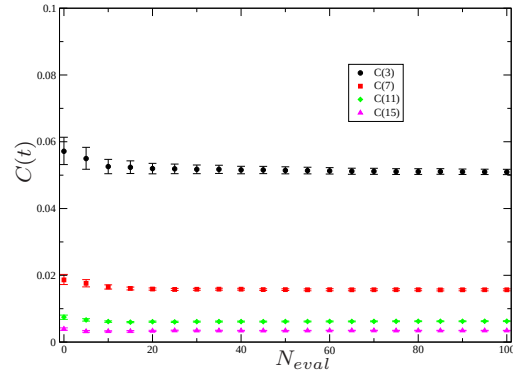
Figure 6.2: The low low pseudoscalar correlation function (red), the high mode parts approximated by a point source (green) and the sum (blue) using eigenvectors of H_{ov} (a) H_{DWF} (b) and $D_{oo}^\dagger D_{oo}$ (c). This is the γ_5 , γ_5 correlation function that has the strongest overlap with the pion. Light quark masses are 0.01. Figure (d) shows the effective mass for each of the LMA methods as well as point sources on the same number of configurations. 100 eigenvectors of each type.



(a) 4d LMA



(b) 5d LMA



(c) 5d preconditioned LMA

Figure 6.3: The low mode averaged correlation function at different time steps as a function of the number of eigenvectors for the three different types of LMA. As we add more eigenvectors, we rapidly saturate the correlation function, especially at late times. The saturation is independent of the operator type.

Firstly in figure 6.2 we show low mode averaging with each of the different types of eigenvectors using 30 configurations on the 0.01, 16^3 ensemble. We have used one hundred eigenvectors in all cases and we have converged the 4d eigenvectors to a precision of 10^{-4} , this will be justified later. The 5d eigenvectors are converged to 10^{-10} . We calculate the low mode correlator C_{LL} and the remainder $C_{HL} + C_{LH} + C_{HH} \approx C^{(point)} - C_{LL}^{(point)}$ as well as their sum $C^{LMA}(t)$ for the pion correlator, $\Gamma_1 = \Gamma_2 = \gamma_5$. As the lightest hadron, the pion should benefit most from low mode averaging, as has been found previously in the literature [47], [72]. We find that the pion correlation function is dominated by its low mode part at late times in all three cases. Figure 6.2(d) reveals two significant facts firstly, as hoped, including 100 eigenvectors dramatically reduces the statistical uncertainty. Secondly, low mode averaging using eigenvectors of the 4d, 5d or 5d preconditioned operator gives the same reduction in statistical error. Though it seemed possible that the 4d operator would have the strongest overlap with the physical low modes making it better suited for LMA, in practice for LMA the choice of operator with a fixed number of eigenvectors doesn't make any difference.

In figures 6.3 we plot $C(t)$ for the pion at selected values of t versus the number of eigenvectors included in the low mode part for each of the three methods. After roughly 20 eigenvectors are included there is a plateau, the error still reduces after this point but more slowly. All three plots look very similar, none plateau earlier and the errors don't reduce any faster depending on the operator. Thus there is no operator that minimises the number of eigenvectors required for LMA to produce a given statistical error.

The cost of the 4d preconditioned operator depends critically on the convergence precision of the inner CG. In figure 6.4 we show how the convergence precision of the eigenvectors affects the accuracy of the mass extracted from the LMA pion correlator. The figure shows m_π with error bars as a function of $-\log(\epsilon)$, so precision increases to the right. Once the eigenvector residual is smaller than the magnitude of the error on the mass then m_π is relatively insensitive to the precision with which the eigenvectors and hence the inner CG are converged. In order to minimize the cost we choose a required residual for our eigenvectors of $\epsilon = 10^{-4}$, which is on the conservative side but seems reasonable if we are aiming to measure quantities to an accuracy of one part in 10^{-3} . The

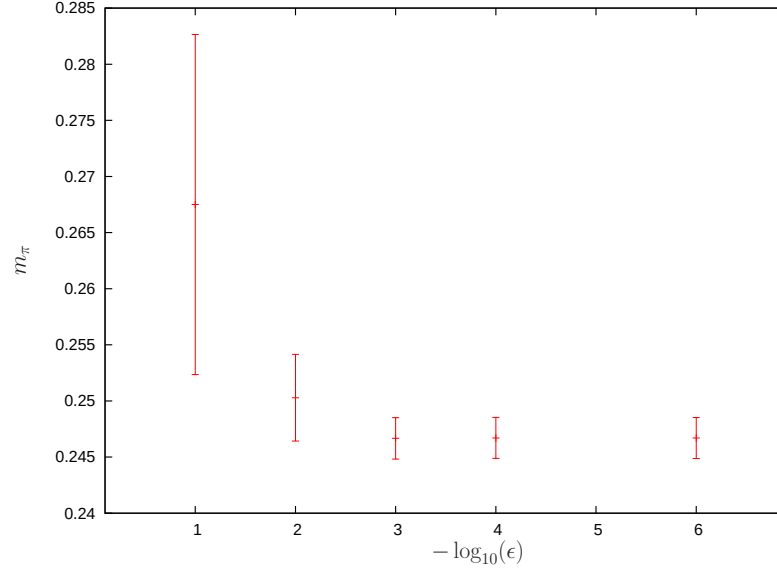


Figure 6.4: The effective mass of the pion using LMA with H_{ov} eigenvectors converged with different precision. Using 30 configurations, quark mass 0.01, fit range $[9 : 15]$. Once the eigenvectors are converged to better than 10^{-3} no further reduction in statistical error is seen.

residual should be of the order of the statistical error but need not be much smaller.

We now compare the pion mass calculated via the three LMA methods and point sources at fixed cost. The cheapest eigenvectors to calculate are those of $D_{oo}^\dagger D_{oo}$. On the thirty configurations used here the average time to calculate 100 eigenvectors, a point source propagator and calculate the low mode correlation function on 128, x86 nodes of the IRIDIS cluster for each of the three methods, as well as the average time to calculate a propagator (in units where $1 = T_{prop} = 1200s$) was,

$$\begin{aligned}
 4d &= 9.27 \\
 5d &= 4.43 \\
 5dprec &= 1.98 \\
 \text{Propagator} &= 1.
 \end{aligned} \tag{6.56}$$

The preconditioned eigenvector calculation is twice as fast as the unpreconditioned and four times as fast as the effective overlap. Since we have found very

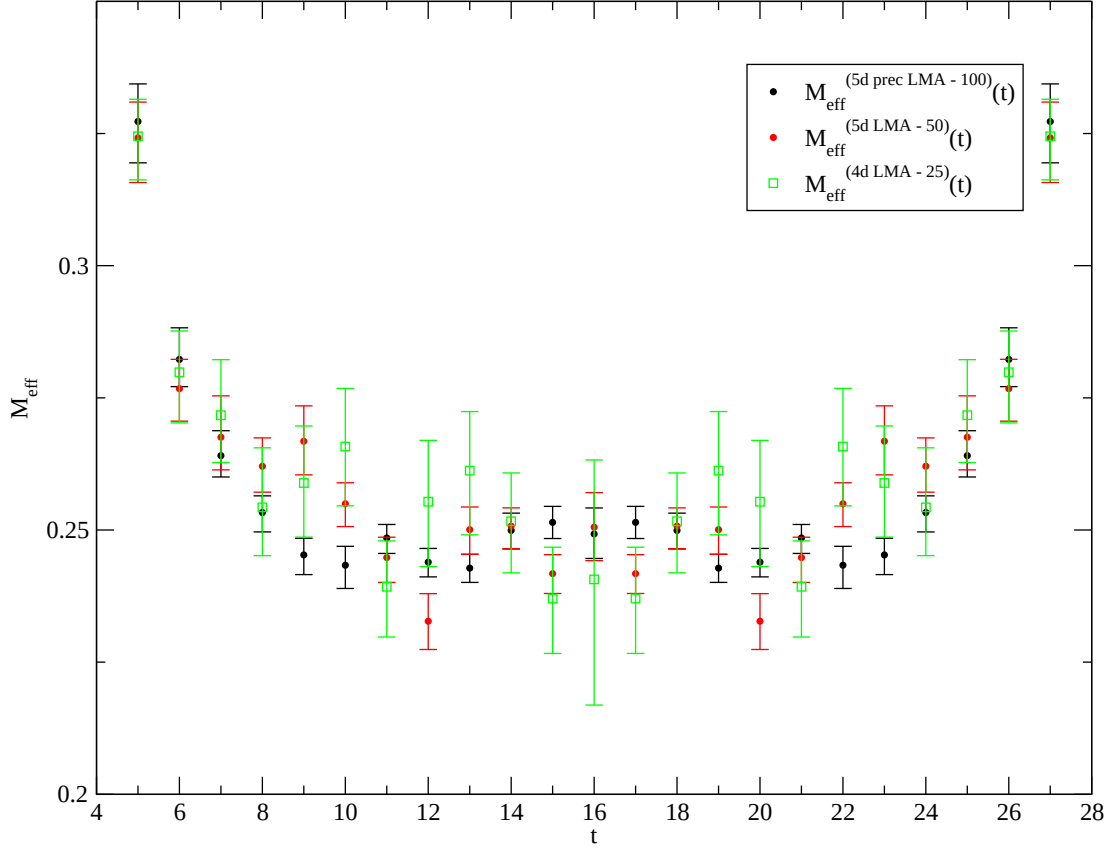


Figure 6.5: The effective mass of the pion for the three different LMA operators at roughly fixed cost. 100 5d preconditioned eigenvectors (red), 50 unpreconditioned (green), 25 4d (black). The fit range for 6.57 is chosen as $[9, 15]$ deep in the region of low mode dominance.

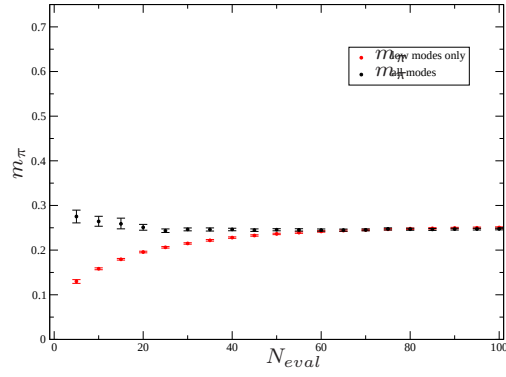
little difference for equivalent numbers of eigenvectors of different operators the preconditioned operator looks to be the best since it can be calculated quickest. This is confirmed in figure 6.5, where we plot the effective masses obtained with each method at roughly fixed cost. The masses are,

$$\begin{aligned}
 5d_{\text{prec}} &= 0.2472 \pm 0.0019 \\
 5d &= 0.2464 \pm 0.0024 \\
 4d &= 0.2475 \pm 0.0039 \\
 \text{Point} &= 0.2567 \pm 0.0045.
 \end{aligned} \tag{6.57}$$

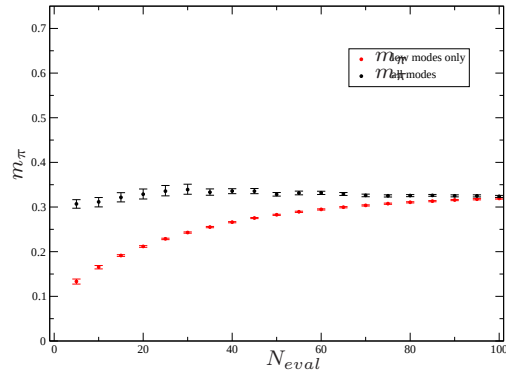
This is to be compared with the pion mass calculated from a much larger number of configurations in reference [2], 0.247 ± 0.003 . We obtain a good answer with good stat error, even though [2] used ≈ 20 times as many configurations. Reference [2] quoted an inflated error because greater than statistical deviations were observed using different source locations. A key advantage of the LMA approach is that all possible source locations are included in an efficient way. In equation 6.57 we also show the equivalent cost point source result (with one source position) calculated by measuring point source correlation functions on 60 configurations instead of 30. The error is reduced by a factor ≈ 2.5 with LMA and the final answer is closer to the correct value of 0.247 and we do not have to account for uncertainties arising from the choice of source position. We conclude that eigenvectors of $D_{oo}^\dagger D_{oo}$ are most suitable for LMA with domain wall fermions. LMA with 100 eigenvectors can be accomplished with a cost equivalent to roughly two point source propagators and the increase in statistical accuracy more than compensates.

6.4.2 Mass Dependence

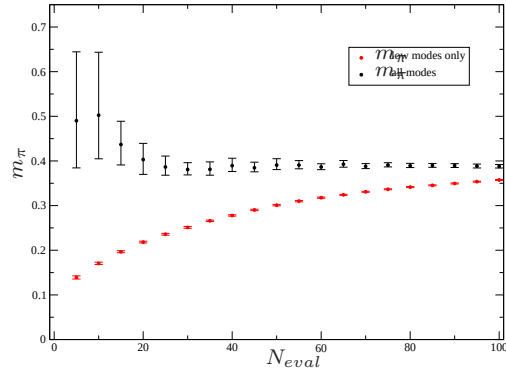
As the physical states depend more on low energy dynamics the low mode part should dominate. Thus as the pion mass decreases the low mode piece should contribute more. In figure 6.6 we show the pion mass calculated by fitting the effective mass $\log \left(\frac{C(t)}{C(t+1)} \right)$ with $C = C_{LL}$ and $C = C^{LMA}$ in the range $t = [9, 15]$ as a function of the number of eigenvectors included in the low mode piece. When $C_{LL} = C^{LMA}$ the high mode piece is not contributing and the two masses will be



(a) 0.01 ensemble



(b) 0.02 ensemble



(c) 0.03 ensemble

Figure 6.6: The pion mass, calculated by fitting $C_{LL}(t)$ (red) and $C^{LMA}(t)$ (black) in the range $[9, 15]$ for different light quark masses, versus the number of eigenvectors. Using 30 gauge configurations in each case. The low modes dominate at lighter pion masses and large time separations.

equal, in this situation we have low mode dominance. Figures 6.6 (a), (b) and (c) show the pion mass as a function of number of eigenvectors on the 0.01, 0.02 and 0.03 mass ensembles respectively. The mass extracted from C_{LL} approaches the mass from the full correlator as the number of eigenvectors, N_{vecs} , increases. For the 0.01 the pion mass extracted from C_{LL} is identical (within errors) to that extracted from the full correlator C^{LMA} for $N_{vecs} > 60$. For 0.02 this happens around $N_{vecs} > 100$ and for 0.03 this will occur somewhere outside the range of our data.

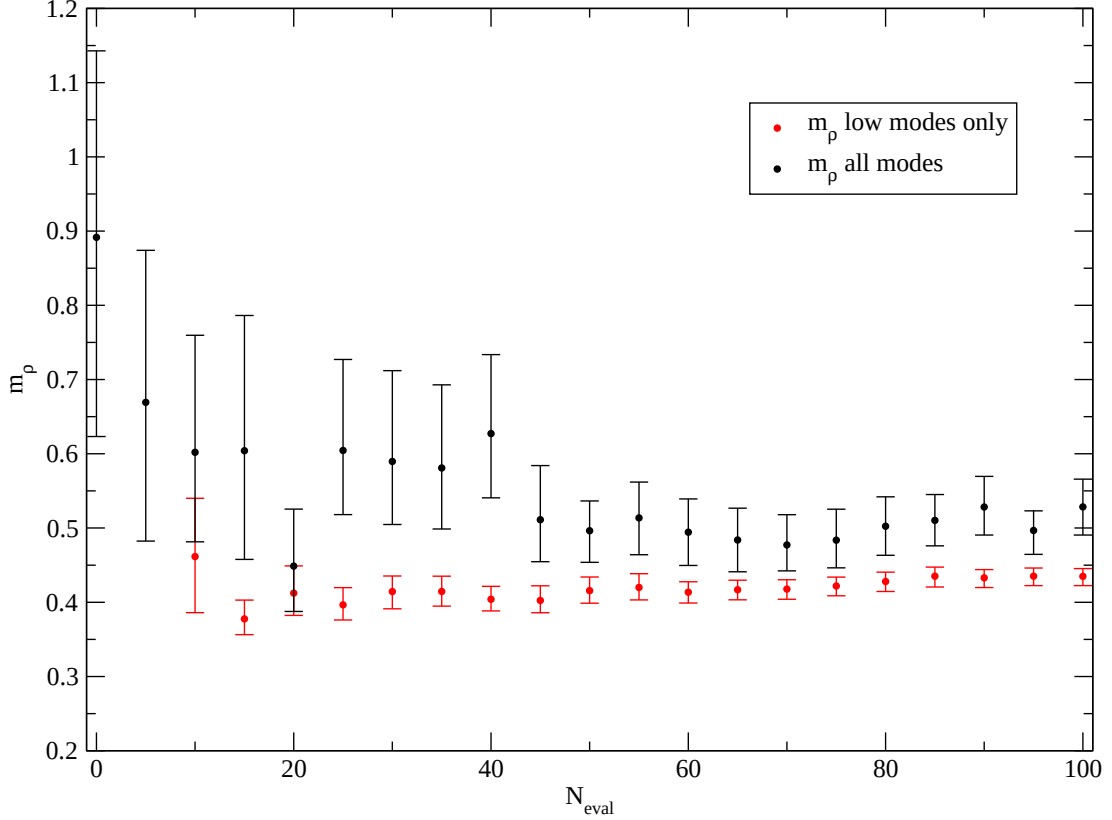


Figure 6.7: The ρ mass calculated by fitting $C_{LL}(t)$ (red) and $C^{LMA}(t)$ (black) in the range $[9, 15]$ versus the number of eigenvectors. Using 30 gauge configurations on the 0.01 ensemble.

The pion is a special case where LMA works extremely well. Other, heavier meson states are not as affected by low energy dynamics and LMA will not be as beneficial. The next heaviest meson to look at is the ρ meson, which is ≈ 5.5 times heavier than the pion. Using the 0.01 ensemble we have made a similar plot to figure 6.6 in figure 6.7 but for the ρ meson. As expected the low modes

contribute less and the gain in statistical accuracy, though significant, is not as impressive. Comparing with reference [2] which has $m_\rho = 0.549(20)$ we get,

$$m_\rho = 0.532(64), \quad (6.58)$$

which agrees within statistical errors and, considering only 30 configurations were used, is a good result.

As the mass decreases LMA is more and more effective for the pion. Physical simulations are approaching the physical pion mass and this indicates that LMA is a good idea for future simulations. However LMA alone is not a panacea and is most beneficial for low energy observables. As lattice simulations approach physical parameters it is expected that LMA will bring greater benefits. For particles of intermediate masses, not too heavy or too light, other techniques should be explored [15] in conjunction with LMA.

6.4.3 η Meson

In calculations of disconnected diagrams low mode averaging is usually hoped to be of benefit. Traces over the spatial lattice like $\sum_{\vec{x}} \text{Tr}[S(x, x)\Gamma]$ are difficult to evaluate. Very many point sources would have to be used to get a good approximation to the sum. More often smearing, to get a better overlap with the physical state, and multiple sources and sinks are used though there is the question of the optimal number of sources and smearing parameters. An alternative is to estimate the traces with stochastic sources, see the recent papers [7], [85] [112] and references therein on estimating the η and η' masses with stochastic sources. The problem with stochastic sources is that we introduce extra noise into an already noisy calculation. In this thesis we try to use LMA in conjunction with stochastic sources to estimate disconnected diagrams, this approach has already met with some success using Wilson fermions [119] and is attempted here with domain wall fermions for the first time. We use low modes of $D_{oo}^\dagger D_{oo}$ throughout this section.

Calculating the η mesons requires the evaluation of three disconnected correlation functions, section 3.1.2. Following reference [39] we require the four

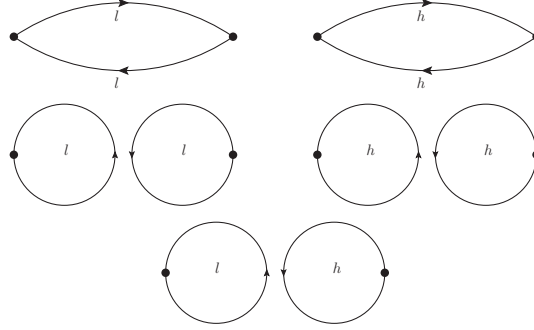


Figure 6.8: The five correlation functions that must be evaluated to calculate the η η' masses. Top left C_{ll} , top right C_{hh} , centre left D_{ll} , centre right D_{hh} , bottom $D_{lh} = D_{hl}$.

Euclidean space correlation functions,

$$C_{\alpha,\beta} = \frac{1}{T} \sum_{\tau} \langle O_{\alpha}(t + \tau)^{\dagger} O_{\beta}(\tau) \rangle \quad (6.59)$$

with $\alpha, \beta \in l, s$ where l is a light quark and s the strange. This matrix can be computed by calculating 5 diagrams, two connected and three disconnected, see figure 6.8.

$$C(t) = \begin{pmatrix} C_{ll} & C_{ls} \\ C_{sl} & C_{ss} \end{pmatrix} = \begin{pmatrix} C_{ll} - 2\mathcal{D}_{ll} & -\sqrt{2}\mathcal{D}_{ls} \\ -\sqrt{2}\mathcal{D}_{sl} & C_{ss} - \mathcal{D}_{ss} \end{pmatrix}. \quad (6.60)$$

and $\mathcal{D}_{ls} = \mathcal{D}_{sl}$. The operators O_l and O_s mix through $\mathcal{D}_{sl}$ so do not create energy eigenstates and the matrix above must be diagonalized.

$$C(t) = A^T D(t) A \quad (6.61)$$

$$D(t) = \begin{pmatrix} e^{-m_{\eta}t} & 0 \\ 0 & e^{-m_{\eta'}t} \end{pmatrix}. \quad (6.62)$$

The eigenvalues are real because $C(t)$ is real and symmetric.

Among the disconnected diagrams $\mathcal{D}_{ll}$, which is also the largest, should benefit from LMA due to the smaller masses involved. As with the connected pieces LMA will be less effective for $\mathcal{D}_{ls}$ and even less so for $\mathcal{D}_{ss}$. However these graphs should be much smaller than $\mathcal{D}_{ll}$. We hope that by accurately measuring the largest

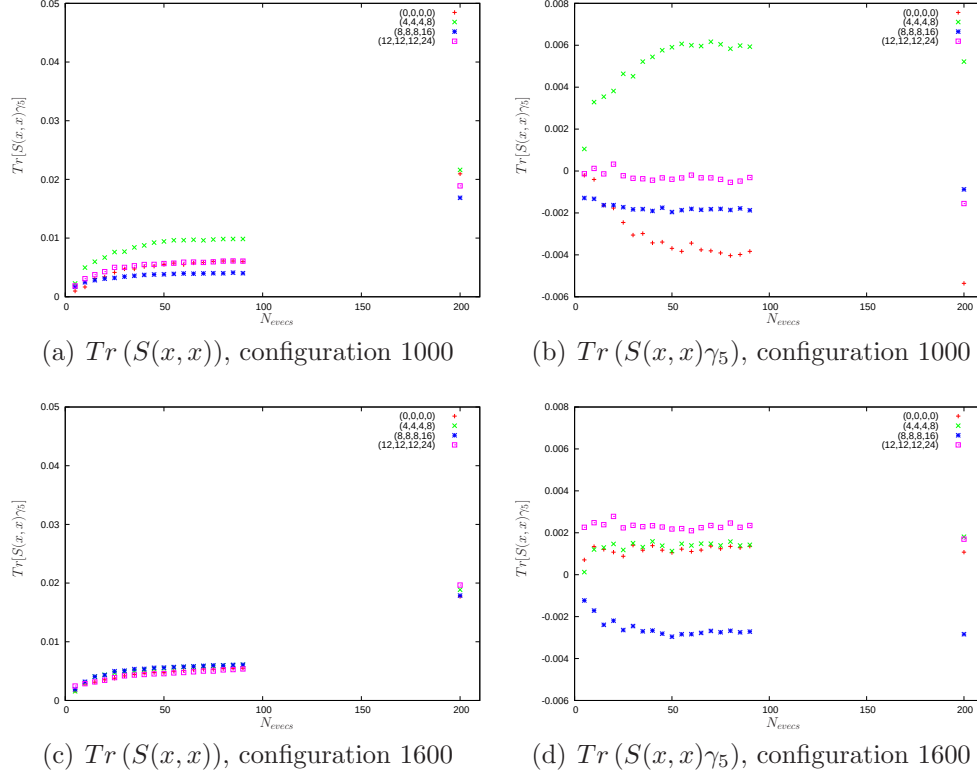


Figure 6.9: Number of modes versus value of the trace. The all mode answer is on the far right. We show 4 different sources on two different configurations of the 16^3 , 0.01 ensemble. $Tr[S(x, x)]$ on the left is not well approximated by its low mode sum, $Tr[S(x, x)\gamma_5]$ is better approximated by the sum of the lowest ≈ 100 eigenvectors.

contribution, $\mathcal{D}_u$, we can obtain accurate measurements of m_η and $m_{\eta'}$. On a real 16^3 configuration computing a quark propagator on every lattice point is prohibitively expensive, we can however calculate $S(x, x)$ for specific points on the lattice, say $x = (0, 0, 0, 0)$. Then traces $Tr[S(x, x)]$ and $Tr[S(x, x)\gamma_5]$ can be evaluated exactly at that point. The low mode decomposition of $S(x, x)$ is also available at that point which we can compare to the exact answer. We plot this in figure 6.9 where we have summed up to 100 eigenmodes to calculate the low mode part of the trace and compare with the exact answer, placed on the far right on the graph. From this plot we see that the trace $Tr[S(x, x)]$ is not well approximated by the lowest modes, the first 100 account for only $\sim 30\%$ of the total. The trace $Tr[S(x, x)\gamma_5]$, which is the one relevant for the η meson, is better approximated by its low mode sum, with the first 100 modes accounting for some $\sim 80\%$ of the total. This plot gives us hope that LMA may be effective for the η .

It is necessary to estimate the high mode pieces, for this we use stochastic methods. Achieving accurate stochastic estimates of correlators depends crucially on the method of dilution, refs [64] and [63], the wrong choice can lead to the signal, which falls exponentially with distance, being swamped by the imperfectly cancelled stochastic noise. For connected correlation functions time dilution is crucial to observe a signal in the noisy background and for our disconnected γ_5 correlators spin dilution is essential; other variables may also be diluted and we explore if this can improve the signal. In order to investigate the question at little cost we study different dilution schemes on the very small $4^3 \times 8$ lattices described in section 5.3. If we do four separate inversions, one for each spin component, we do the spin trace explicitly while doing the other traces stochastically. This is spin dilution. Similarly we can do colour dilution or time dilution by doing an inversion for each spin or time component. We can also do a partial dilution in the time direction, relying on the exponential decay of the propagator to suppress contamination of the signal with noise. For this we split the source into a sum of t_{sep} sources that are non-zero only on timeslices separated by Δt , we call this a comb.

Figure 6.10 shows the correlator

$$D_u(t) = \frac{1}{VT} \sum_{\vec{x}, \vec{y}, \tau} Tr[S_l(x, x)\gamma_5] Tr[S_l(y, y)\gamma_5] \quad (6.63)$$

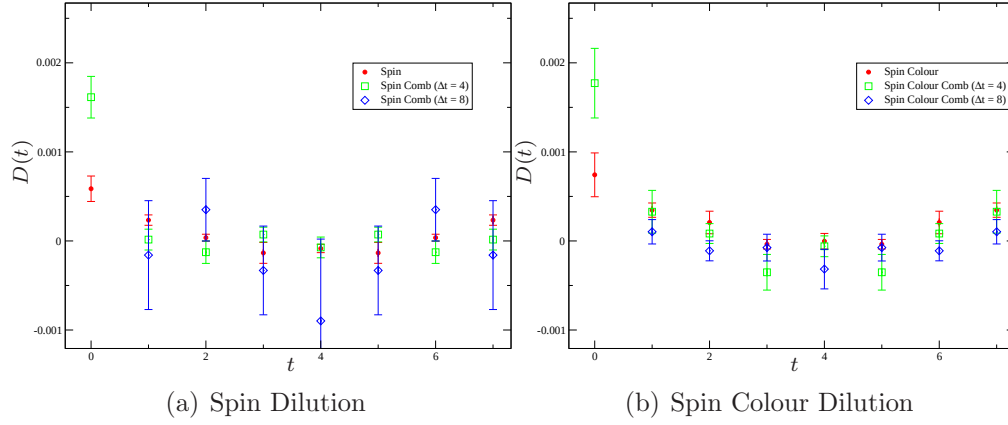


Figure 6.10: Measuring $D(t)$ using 44 configurations on a $4^3 \times 8$ lattice with different dilution schemes. All the data is computed at equal cost with a total of 96 solves of $Dx = b$ for each source type per configuration. The methods without time dilution are most effective.

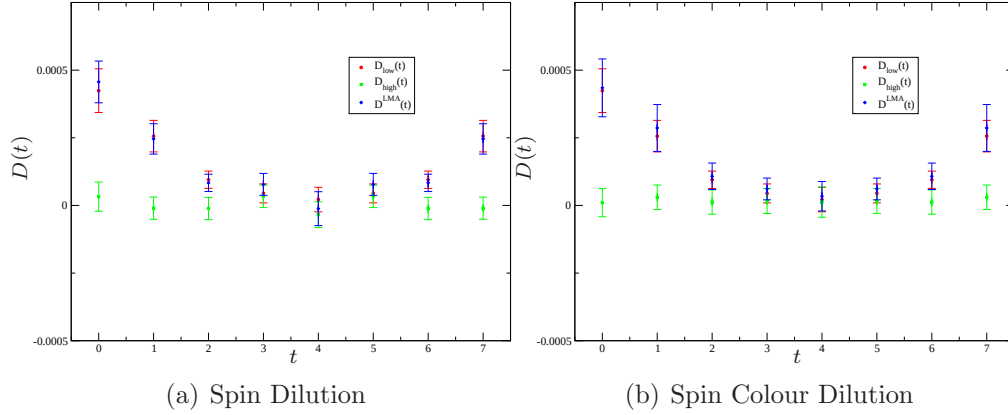


Figure 6.11: $D(t)$ using 44 configurations on a $4^3 \times 8$ lattice with different spin (a) and spin+colour(b) dilution and 100 eigenvectors of the $5d$ preconditioned operator. $D(t)$ is split into low (red), high (blue) and their sum (blue). Spin and spin-colour dilution are almost equivalent with slightly smaller error bars for spin dilution.

at the lightest quark mass evaluated with different stochastic estimators. Figure 6.10 (a) shows spin and spin+comb dilution and figure 6.10(b) shows spin+colour and spin+colour+comb dilution. These results imply that time dilution does not help, in contrast to the connected case where it is essential. The best choices are spin or spin+colour. We plot the low mode part, high mode part and full LMA of the correlator in figure 6.11. From this we can see that with spin diluted sources we have a slightly smaller statistical error on each data point. Figure 6.9 implies that the correlation function should be low mode dominated. The high mode part is a very noisy estimate of a numerically small quantity and for spin sources we have the most noise vectors at fixed cost. These results imply that the best approach is LMA with dilution of the spin component only and this is what we will do for our study of the η meson on $16^3 \times 32$ lattices.

6.4.4 Results

We used the same configurations as reference [39] which are similar to the 16^3 lattices from before: $16^3 \times 32$, Iwasaki gauge, $L_s = 16$ dynamical domain wall fermions. The light quark mass is 0.01 and the strange quark mass is in this case equal to 0.032, much closer to its physical value. This is a test only of the efficacy of LMA for disconnected diagrams for the iso-singlet, we do not attempt a chiral limit or an analysis of the mixing angles involved with the η, η' system. We are only interested in the statistical accuracy with which we can obtain m_η and $m_{\eta'}$ at finite quark mass.

We plot the correlators required for the matrix 6.60 in figure 6.12 with the connected pieces in red and the disconnected in blue. The connected pieces have also been evaluated using LMA with stochastic sampling, however, because they are so much more precise than the disconnected pieces, we can estimate them using only one, time diluted, noise vector. Since we only need the γ_5, γ_5 correlator we can calculate the correlation function with only one Dirac matrix inversion, the *Z2PSWall* method of reference [32]. In figure 6.12 we plot the five correlators of equation 6.60. They have the expected sizes, with $D_{ss} < D_{ls} < D_{ll}$. The disconnected contribution D_{ll} is as large as C_{ll} at late times. This is to be compared with figure 2. of reference [39] where a very similar plot is obtained.

Similarly to the meson, figure 6.3, we plot the disconnected correlation functions at selected times ($t = \{2, 7\}$) as a function of the number of eigenvectors

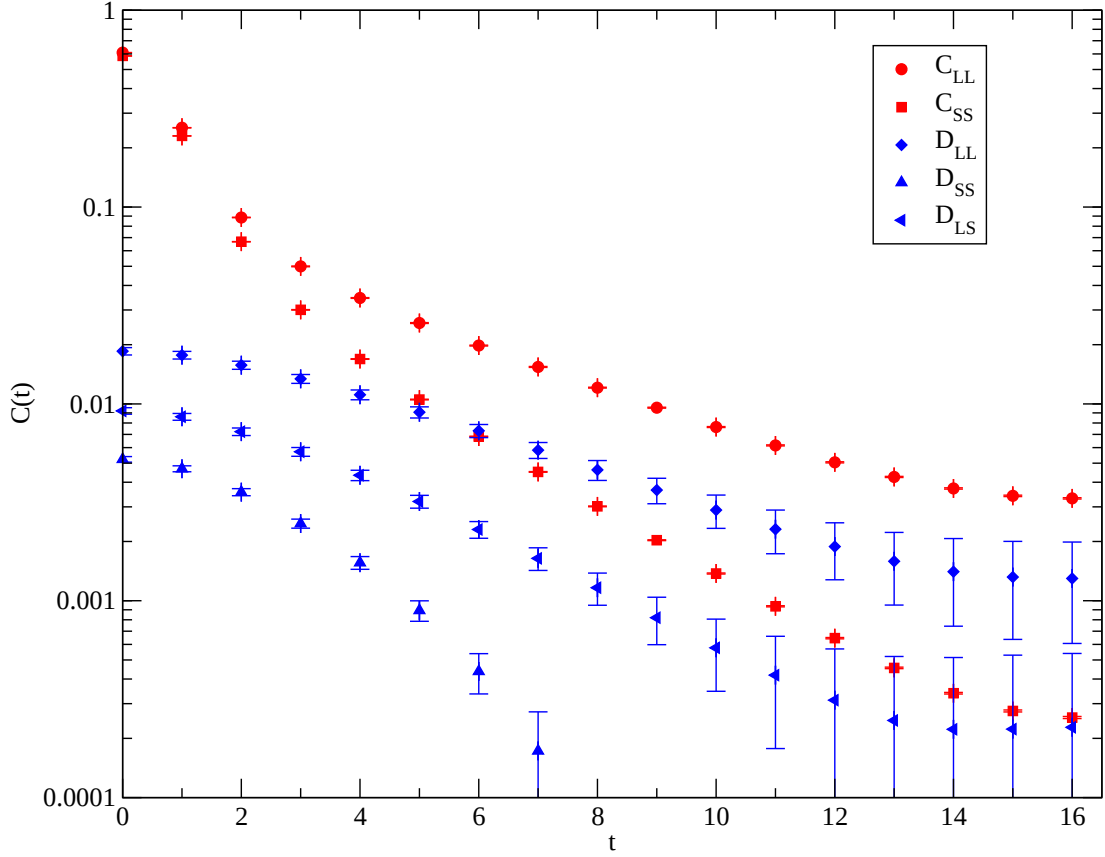


Figure 6.12: The five correlation functions necessary for the η η' masses. Computed on 300 configurations, with 64 spin diluted noise sources for the disconnected diagrams and one *Z2PSWall* source for the connected piece. Using LMA with 100 low modes for each correlation function.

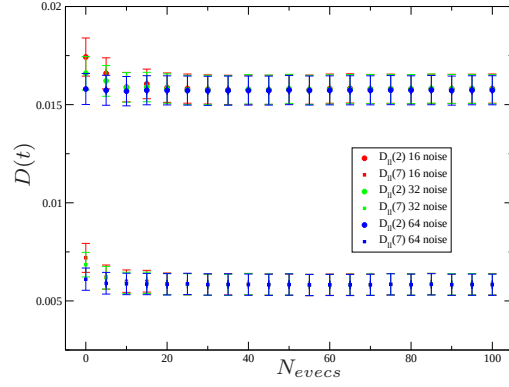
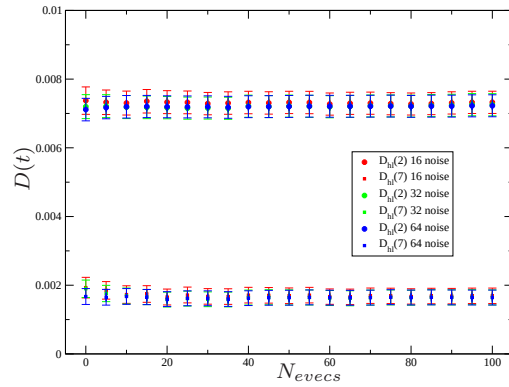
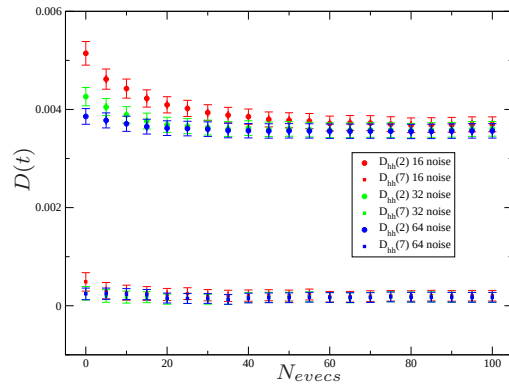
(a) D_{0l} (b) D_{hl} (c) D_{hh}

Figure 6.13: The disconnected correlation functions evaluated with different numbers of noise vectors as a function of the number of eigenvectors used for LMA. At late times and reasonably large N_{evecs} the final value of $D(t)$ is independent of the number of noise vectors.

in figure 6.13. We use different numbers of noise vectors, corresponding to $\{16, 32, 64\}$ Dirac inversions, to estimate the high mode part. We see that for small N_{evect} the number of noise samples is important however by $N_{evect} = 100$ the final point is not very sensitive to the number of noise samples. Even for the D_{hh} correlator, which is least dominated by its low modes, by the time around 60 eigenvectors have been included the final result is insensitive to the number of noise vectors used to estimate the high mode part. At later times, when heavier modes have decayed, this holds even more so. This is taken as a signal that by $N_{evect} \sim 60$ we have low mode dominance. Essentially, with 100 low modes we are doing as well as we can do, the error from the high mode part is not a factor and we are dominated by configuration noise.

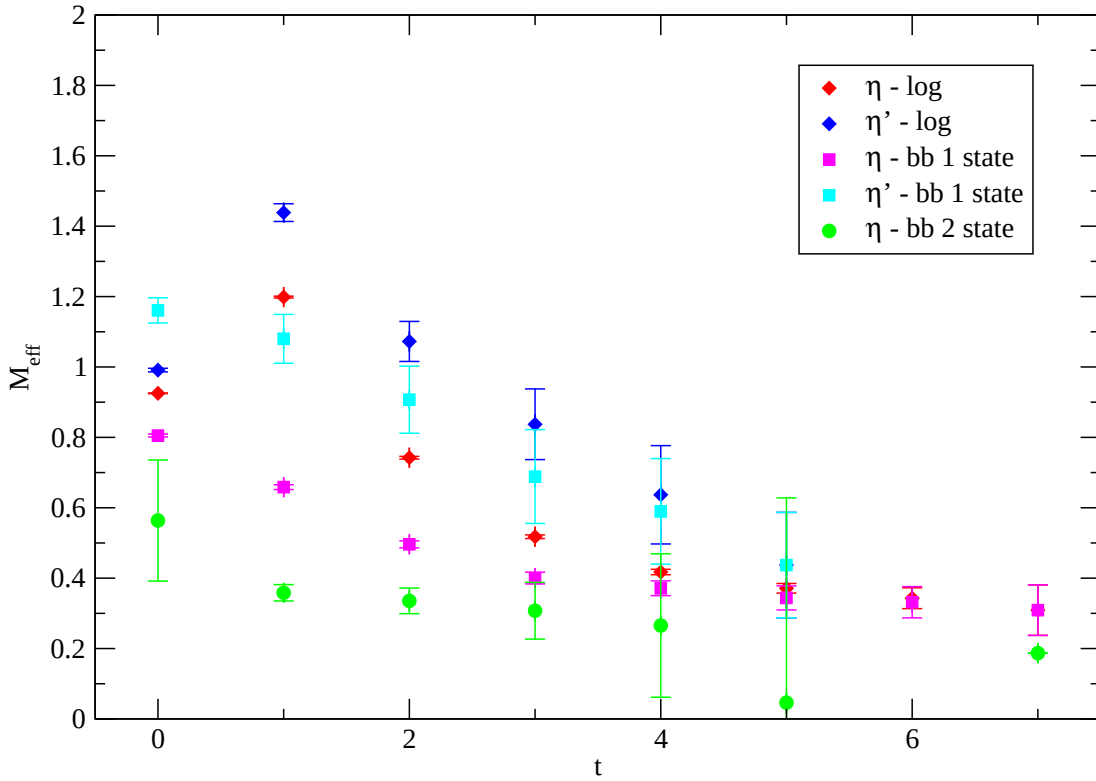


Figure 6.14: The effective mass of the $\eta - \eta'$ mesons as a function of t . Using the log definition and the black box method. The η shows a definite plateau the η' is less clear.

We use the same fit procedure as reference [39] which is to calculate,

$$C(t_0)^{-1}C(t) = A^{-1}D(t - t_0)A \quad (6.64)$$

by diagonalizing the matrix 6.60 to get the eigenvalues $\lambda_1 = e^{m_\eta(t-t_0)}$ and $\lambda_2 = e^{m_{\eta'}(t-t_0)}$. Using $t_0 = 2$ we plot the effective mass in figure 6.14 by two methods: the ratio $\log\left(\frac{\lambda(t)}{\lambda(t+1)}\right)$ and the linear prediction, black box method of section 3.4.1. For the black box method we include all data from t to t_{max} and calculate the effective mass. We assume a single exponential for the η' and $t_{max} = 7$. For the η we put $t_{max} = 9$ and also try a two excited state fit. We consider the η meson first. The effective mass plots show a plateau for each of the three ways of calculating m_{eff} . The black box method is particularly succesful at obtaining an early plateau. Fitting each one we get,

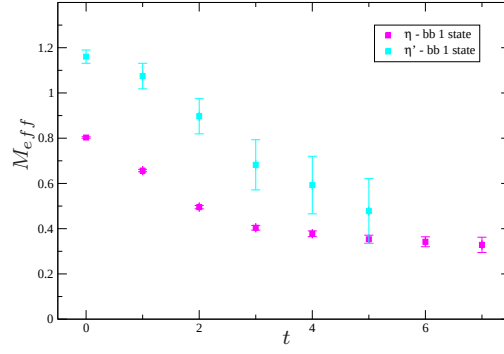
$$\begin{aligned} m_\eta &= 0.3671(130) : \log, \text{ range } [5,7] \\ m_\eta &= 0.3629(176) : \text{bb 1 state}, \text{ range } [4,7] \\ m_\eta &= 0.3547(190) : \text{bb 2 state}, \text{ range } [1,4] \end{aligned} \tag{6.65}$$

where the fit ranges for each method are specified above. The best estimate is from the log definition and has a 3.5% error. For the η' we do not see an obvious plateau, we fit the last two or three data-points as a best estimate and obtain,

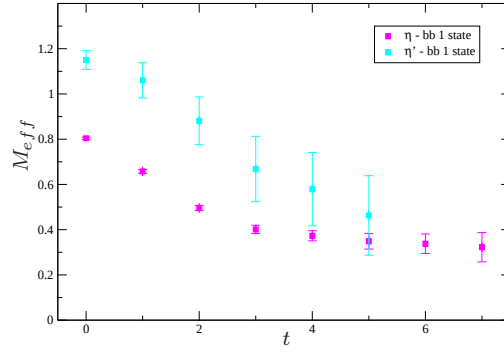
$$\begin{aligned} m_{\eta'} &= 0.568(122) : \log, \text{ range } [4,5] \\ m_{\eta'} &= 0.600(121) : \text{bb 1 state}, \text{ range } [3,5]. \end{aligned} \tag{6.66}$$

The best estimate is from the black box fit and has an error of 20%.

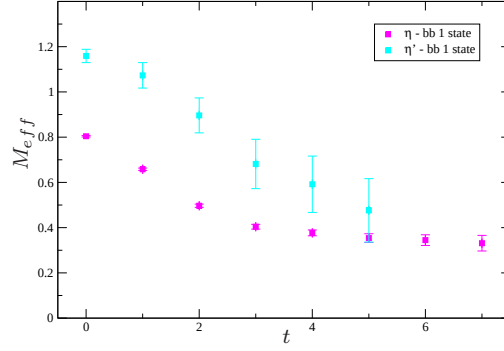
The error is dominated by configuration noise. To see this we reduce the number of eigenvectors to 35 and the number of noise vectors to 16 in figure 6.15(a) we also use 100 eigenvectors and 64 noise vectors and half the number of configurations in figure 6.15(b). Comparing to figure 6.15(c) we see that changing the sampling per configuration has little effect whereas including more configurations has made a significant difference. Fitting the same range as before



(a) 16 noise, 35 eig, 300 configs



(b) 64 noise, 100 eig, 150 configs



(c) 64 noise, 100 eig, 300 configs

Figure 6.15: The black box effective masses, calculated by reducing sampling on each configuration (a) versus sampling fewer configurations (b), to be compared with the result from all samples and configurations (c). The error bars in (b) are quite a bit larger than those in (a) and (c) which are nearly identical, showing we are not limited by sampling on each configuration.

the η and η' masses change to,

$$\begin{aligned}
m_\eta &= 0.3671(130) : 64 \text{ noise, 100 eigenvectors, 300 configs} & (6.67) \\
m_\eta &= 0.3675(132) : 16 \text{ noise, 35 eigenvectors, 300 configs} \\
m_\eta &= 0.3643(214) : 64 \text{ noise, 100 eigenvectors, 150 configs} \\
m_{\eta'} &= 0.600(121) : 64 \text{ noise, 100 eigenvectors, 300 configs} \\
m_{\eta'} &= 0.602(123) : 16 \text{ noise, 35 eigenvectors, 300 configs} \\
m_{\eta'} &= 0.583(157) : 64 \text{ noise, 100 eigenvectors, 150 configs} .
\end{aligned}$$

This implies that we can do a much cheaper calculation with only 16 light quark and 16 heavy quark inversions with 35 eigenvectors per configuration and achieve the same level of accuracy.

6.5 Conclusions

We have formulated LMA with eigenvectors of the domain wall operator using three different approaches. The most cost effective was found to be LMA with low modes of $D_{oo}^\dagger D_{oo}$. For the pion and the rho meson we found that LMA was very effective at reducing the statistical error and grows more effective as the quark masses are reduced. We also calculated the η and η' masses using stochastic sampling together with LMA.

Compared to reference [39], which calculated the η and η' mass on the same number of configurations and obtained an error of 0.7% and 7.5% respectively we have not been able to obtain similarly accurate results using noise sources together with low modes. The Coulomb gauge fixed wall sources of [39] have better overlap with the physical η meson state. The plateaus there begin earlier and the later time data is much cleaner. The advantage of our method is that it is significantly cheaper, using as few as 35 eigenvectors and 16 noise sources we obtain results of the same quality as using 100 eigenvectors and 64 noise sources. We can compare cost using figure 5.11 and 5.10 in chapter 5. Using time units where the cost of an undeflated propagator is 1, calculating 35 eigenvectors costs roughly 0.7, this is a cost shared by [39] which uses the low modes for deflation only. Then [39] does $N_t = 32$ deflated light and heavy propagator solves. Referring to figure 5.13, showing the very weak mass dependence of calculating eigenvectors and deflated

propagator solves the cost is roughly $(2 \times 0.7) + (2 \times 0.4 \times 32) = 27$. Our method reuses the eigenvectors for LMA and requires $(2 \times 0.7) + (2 \times 0.4 \times \frac{16+1}{12}) \sim 2.5$ for propagator inversions. Thus our method is some 10 times cheaper. Though the final values have not been obtained as accurately as hoped the much reduced cost still makes this method viable, especially if we can incorporate the good overlap of wall sources with the meson wavefunction. On future (and present!) lattices with larger N_t and smaller quark mass the cost of doing a propagator inversion on every timeslice will grow significantly whereas the cost of our method grows more slowly, so on larger lattices at smaller quark masses our method may be the superior one. The question can only be answered by simulations, which we leave for future work.

Bibliography

- [1] C. Allton et al. *Phys. Rev.*, D76:014504, 2007.
- [2] C. Allton et al. Physical Results from 2+1 Flavor Domain Wall QCD and SU(2) Chiral Perturbation Theory. *Phys. Rev.*, D78:114509, 2008.
- [3] Leandro G. Almeida and Christian Sturm. Two-loop matching factors for light quark masses and three-loop mass anomalous dimensions in the RI/SMOM schemes. *Phys.Rev.*, D82:054017, 2010.
- [4] Claude Amsler et al. Review of Particle Physics. *Phys.Lett.*, B667:1–1340, 2008.
- [5] David J. Antonio et al. Localization and chiral symmetry in three flavor domain wall QCD. *Phys.Rev.*, D77:014509, 2008.
- [6] D.J. Antonio et al. First results from 2+1 Flavor Domain Wall QCD: Mass Spectrum, Topology Change and Chiral Symmetry with $L(s) = 8$. *Phys.Rev.*, D75:114501, 2007.
- [7] S. Aoki et al. An estimate of the eta and eta' meson masses in $N(f) = 2+1$ lattice QCD. *PoS*, LAT2006:204, 2006.
- [8] Y. Aoki, P.A. Boyle, N.H. Christ, C. Dawson, M.A. Donnellan, et al. Non-perturbative renormalization of quark bilinear operators and B(K) using domain wall fermions. *Phys.Rev.*, D78:054510, 2008.
- [9] Y. Aoki et al. Continuum Limit of B_K from 2+1 Flavor Domain Wall QCD. *Phys. Rev.*, D84:014503, 2011.
- [10] Y. Aoki et al. Continuum Limit Physics from 2+1 Flavor Domain Wall QCD. *Phys.Rev.*, D83:074508, 2011.

- [11] M. Atiyah and I. M. Singer. The Index of Elliptic Operators. *The Annals of Mathematics*, 93, 1971.
- [12] Taegil Bae et al. Kaon B -parameter from improved staggered fermions in $N_f = 2 + 1$ QCD. 2011.
- [13] Gunnar Bali, Luca Castagnini, and Sara Collins. Meson and baryon masses with low mode averaging. *PoS, LATTICE2010:096*, 2010.
- [14] Gunnar S. Bali and Peter Boyle. Perturbative Wilson loops with massive sea quarks on the lattice. 2002.
- [15] Gunnar S. Bali, Sara Collins, and Andreas Schafer. Effective noise reduction techniques for disconnected loops in Lattice QCD. *Comput. Phys. Commun.*, 181:1570–1583, 2010.
- [16] D. Becirevic, Philippe Boucaud, J.P. Leroy, J. Micheli, O. Pene, et al. Asymptotic behavior of the gluon propagator from lattice QCD. *Phys.Rev.*, D60:094509, 1999.
- [17] Paulo F. Bedaque. Aharonov-Bohm effect and nucleon nucleon phase shifts on the lattice. *Phys.Lett.*, B593:82–88, 2004.
- [18] A.A. Belavin, Alexander M. Polyakov, A.S. Schwartz, and Yu.S. Tyupkin. Pseudoparticle Solutions of the Yang-Mills Equations. *Phys.Lett.*, B59:85–87, 1975.
- [19] G. Bergner and J. Wuilloud. Acceleration of the Arnoldi method and real eigenvalues of the non-Hermitian Wilson-Dirac operator. *Comput. Phys. Commun.*, 183:299–304, 2012.
- [20] Wolfgang Bietenholz, R. Brower, S. Chandrasekharan, and U.J. Wiese. Progress on perfect lattice actions for QCD. *Nucl.Phys.Proc.Suppl.*, 53:921–934, 1997.
- [21] Erek Bilgici, Antonino Flachi, Etsuko Itou, Masafumi Kurachi, C.-J David Lin, et al. A New scheme for the running coupling constant in gauge theories using Wilson loops. *Phys.Rev.*, D80:034507, 2009.

- [22] Sundance O. Bilson-Thompson, Derek B. Leinweber, and Anthony G. Williams. Highly improved lattice field strength tensor. *Annals Phys.*, 304:1–21, 2003.
- [23] T. Blum, P. Chen, Norman H. Christ, C. Cristian, C. Dawson, et al. Quenched lattice QCD with domain wall fermions and the chiral limit. *Phys.Rev.*, D69:074502, 2004.
- [24] T. Blum, Norman H. Christ, C. Cristian, C. Dawson, George Tamminga Fleming, et al. Nonperturbative renormalization of domain wall fermions: Quark bilinears. *Phys.Rev.*, D66:014504, 2002.
- [25] Achim Bode et al. First results on the running coupling in QCD with two massless flavors. *Phys.Lett.*, B515:49–56, 2001.
- [26] Frederic D.R. Bonnet, Derek B. Leinweber, Anthony G. Williams, and James M. Zanotti. Improved smoothing algorithms for lattice gauge theory. *Phys.Rev.*, D65:114510, 2002.
- [27] Artan Borici. Truncated overlap fermions: The Link between overlap and domain wall fermions. pages 41–52, 1999.
- [28] S. Borsanyi, S. Durr, Z. Fodor, C. Hoelbling, S.D. Katz, et al. High-precision scale setting in lattice QCD. 2012.
- [29] P. A. Boyle. The BAGEL assembler generator. *Computer Physics Communications*, 180/12:2739, 2009.
- [30] P.A. Boyle. Four momentum boosted fermion fields. *Nucl.Phys.Proc.Suppl.*, 129:358–360, 2004.
- [31] P.A. Boyle, J.M. Flynn, A. Juttner, C.T. Sachrajda, and J.M. Zanotti. Hadronic form factors in Lattice QCD at small and vanishing momentum transfer. *JHEP*, 0705:016, 2007.
- [32] P.A. Boyle, A. Juttner, C. Kelly, and R.D. Kenway. Use of stochastic sources for the lattice determination of light quark physics. *JHEP*, 0808:086, 2008.

- [33] P.A. Boyle, A. Juttner, R.D. Kenway, C.T. Sachrajda, S. Sasaki, et al. K(l3) semileptonic form-factor from 2+1 flavour lattice QCD. *Phys.Rev.Lett.*, 100:141601, 2008.
- [34] Jr. Callan, Curtis G. and Jeffrey A. Harvey. Anomalies and Fermion Zero Modes on Strings and Domain Walls. *Nucl.Phys.*, B250:427, 1985.
- [35] D. Calvetti, L. Reichel, and D. C. Sorensen. An implicitly restarted lanczos method for large symmetric. . . *ETNA*, 2:1–21, 1994.
- [36] P. Chen, Norman H. Christ, George Tamminga Fleming, A. Kaehler, C. Malureanu, et al. Domain wall fermion zero modes on classical topological backgrounds. *Phys.Rev.*, D59:054508, 1999.
- [37] T.P. Cheng and L.F. Li. GAUGE THEORY OF ELEMENTARY PARTICLE PHYSICS. 1985.
- [38] K.G. Chetyrkin and A. Retey. Renormalization and running of quark mass and field in the regularization invariant and MS-bar schemes at three loops and four loops. *Nucl.Phys.*, B583:3–34, 2000.
- [39] N. H. Christ et al. The η and η' mesons from Lattice QCD. *Phys. Rev. Lett.*, 105:241601, 2010.
- [40] M.A. Clark and A.D. Kennedy. The RHMC algorithm for two flavors of dynamical staggered fermions. *Nucl.Phys.Proc.Suppl.*, 129:850–852, 2004.
- [41] M.A. Clark, A.D. Kennedy, and Z. Sroczynski. Exact 2+1 flavour RHMC simulations. *Nucl.Phys.Proc.Suppl.*, 140:835–837, 2005.
- [42] Michael Creutz. Confinement, chiral symmetry, and the lattice. 2011.
- [43] C.T.H. Davies et al. High precision lattice QCD confronts experiment. *Phys.Rev.Lett.*, 92:022001, 2004.
- [44] G.M. de Divitiis, R. Petronzio, and N. Tantalo. On the discretization of physical momenta in lattice QCD. *Phys.Lett.*, B595:408–413, 2004.
- [45] P. de Forcrand et al. Renormalization group flow of SU(3) lattice gauge theory: Numerical studies in a two coupling space. *Nucl.Phys.*, B577:263–278, 2000.

- [46] Philippe de Forcrand, Margarita Garcia Perez, and Ion-Olimpiu Stamatescu. Topology of the $SU(2)$ vacuum: A Lattice study using improved cooling. *Nucl.Phys.*, B499:409–449, 1997.
- [47] Thomas DeGrand and Stefan Schaefer. Improving meson two-point functions in lattice qcd. *COMPUT.PHYS.COMMUN.*, 159:185, 2004.
- [48] Bertrand Delamotte. An Introduction to the nonperturbative renormalization group. 2007.
- [49] Shao-Jing Dong and Keh-Fei Liu. Stochastic estimation with $Z(2)$ noise. *Phys.Lett.*, B328:130–136, 1994.
- [50] A. Donini, V. Gimenez, G. Martinelli, M. Talevi, and A. Vladikas. Nonperturbative renormalization of lattice four fermion operators without power subtractions. *Eur.Phys.J.*, C10:121–142, 1999.
- [51] A. Donini, G. Martinelli, Christopher T. Sachrajda, M. Talevi, and A. Vladikas. Nonperturbative renormalization of the lattice Delta $s = 2$ four fermion operator. *Phys.Lett.*, B360:83–92, 1995.
- [52] Nikos Drakos and Ross Moore. Jackknife Error Estimates. 2002.
- [53] D.S.Watkins. Understanding the QR algorithm II.
- [54] S. Duane, A.D. Kennedy, B.J. Pendleton, and D. Roweth. Hybrid Monte Carlo. *Phys.Lett.*, B195:216–222, 1987.
- [55] S. Durr et al. Precision computation of the kaon bag parameter. *Phys.Lett.*, B705:477–481, 2011.
- [56] Robert G. Edwards and Urs M. Heller. Domain wall fermions with exact chiral symmetry. *Phys.Rev.*, D63:094505, 2001.
- [57] Robert G. Edwards, Urs M. Heller, and Rajamani Narayanan. The Hermitian Wilson-Dirac operator in smooth $SU(2)$ instanton backgrounds. *Nucl.Phys.*, B522:285–297, 1998.
- [58] Robert G. Edwards, Urs M. Heller, and Rajamani Narayanan. A study of chiral symmetry in quenched QCD using the overlap Dirac operator. *Phys.Rev.*, D59:094510, 1999.

-
- [59] S. Elitzur. Impossibility of Spontaneously Breaking Local Symmetries. *Phys.Rev.*, D12:3978–3982, 1975.
- [60] George Tamminga Fleming. Finite temperature QCD with domain wall fermions. 2001. Ph.D. Thesis hep-lat/0404020.
- [61] George Tamminga Fleming. What can lattice QCD theorists learn from NMR spectroscopists? *JLAB-THY-04-13*, pages 143–152, 2004.
- [62] J.M. Flynn, A. Juttner, and C.T. Sachrajda. A Numerical study of partially twisted boundary conditions. *Phys.Lett.*, B632:313–318, 2006.
- [63] Justin Foley, K. Jimmy Juge, Alan O’Cais, Mike Peardon, Sinead M. Ryan, et al. Practical all-to-all propagators for lattice QCD. *Comput.Phys.Commun.*, 172:145–162, 2005.
- [64] M. Foster and Christopher Michael. Quark mass dependence of hadron masses from lattice QCD. *Phys.Rev.*, D59:074503, 1999.
- [65] J. G. F. Francis. The QR Transformation A Unitary Analogue to the LR Transformation-Part 1. *The Computer Journal*, 4 (3):265–271, 1961.
- [66] J. G. F. Francis. The QR Transformation-Part 2. *The Computer Journal*, 4 (4):332–345, 1962.
- [67] Kazuo Fujikawa. Path Integral Measure for Gauge Invariant Fermion Theories. *Phys. Rev. Lett.*, 42:1195, 1979.
- [68] Vadim Furman and Yigal Shamir. Axial symmetries in lattice QCD with Kaplan fermions. *Nucl.Phys.*, B439:54–78, 1995.
- [69] Christof Gattringer, Meinulf Gockeler, C.B. Lang, Paul E.L. Rakow, and Andreas Schafer. Comparing lattice Dirac operators in smooth instanton backgrounds. *Phys.Lett.*, B522:194–200, 2001.
- [70] Christof Gattringer and Christian B. Lang. Quantum chromodynamics on the lattice. *Lect.Notes Phys.*, 788:1–211, 2010.
- [71] Paul H. Ginsparg and Kenneth G. Wilson. A Remnant of Chiral Symmetry on the Lattice. *Phys.Rev.*, D25:2649, 1982.

-
- [72] Leonardo Giusti, P. Hernandez, M. Laine, P. Weisz, and H. Wittig. Low-energy couplings of QCD from current correlators near the chiral limit. *JHEP*, 04:013, 2004.
- [73] Leonardo Giusti, C. Hoelbling, M. Luscher, and H. Wittig. Numerical techniques for lattice QCD in the epsilon regime. *Comput.Phys.Commun.*, 153:31–51, 2003.
- [74] M. Gockeler, R. Horsley, H. Oelrich, H. Perlt, D. Petters, et al. Nonperturbative renormalization of composite operators in lattice QCD. *Nucl.Phys.*, B544:699–733, 1999.
- [75] Gene H. Golub and Charles F. van Van Loan. *Matrix Computations*. 1996.
- [76] Martin Gorbahn and Sebastian Jager. Precise $\overline{\text{MS}}$ light-quark masses from lattice QCD in the RI/SMOM scheme. *Phys.Rev.*, D82:114001, 2010.
- [77] J. A. Gracey. Three loop anomalous dimension of non-singlet quark currents in the RI' scheme. *Nucl. Phys.*, B662:247–278, 2003.
- [78] J.A. Gracey. RI'/SMOM scheme amplitudes for quark currents at two loops. *Eur.Phys.J.*, C71:1567, 2011.
- [79] D. J. Gross and Frank Wilczek. ULTRAVIOLET BEHAVIOR OF NON-ABELIAN GAUGE THEORIES. *Phys. Rev. Lett.*, 30:1343–1346, 1973.
- [80] A. Hasenfratz, R. Hoffmann, and F. Knechtli. The Static potential with hypercubic blocking. *Nucl.Phys.Proc.Suppl.*, 106:418–420, 2002.
- [81] M. R. Hestenes and E. Stiefel. Methods of Conjugate Gradients for Solving Linear Systems. *J.Res.Nat.But.Stand.*, 49:409–436, 1952.
- [82] Timothy J. Hollowood. 6 Lectures on QFT, RG and SUSY. 2009.
- [83] Y. Iwasaki. RENORMALIZATION GROUP ANALYSIS OF LATTICE THEORIES AND IMPROVED LATTICE ACTION. 2. FOUR-DIMENSIONAL NONABELIAN SU(N) GAUGE MODEL. 1983.
- [84] Taku Izubuchi. Extended Domain-Wall Fermion operator , its equivalence to the Twisted-mass Wilson Fermions, and its numerical applications. 2011.

-
- [85] K. Jansen, Christopher Michael, and C. Urbach. The eta' meson from lattice QCD. *Eur. Phys. J.*, C58:261–269, 2008.
- [86] Karl Jansen. Domain wall fermions and chiral gauge theories. *Phys.Rept.*, 273:1–54, 1996.
- [87] Chris Johnson and A.D. Kennedy. Bounds on the convergence of Ritz values from Krylov subspaces to interior eigenvalues of Hermitean matrices. 2011.
- [88] Andreas Juttner. Review: The FLAG working group. 2011.
- [89] David B. Kaplan. A Method for simulating chiral fermions on the lattice. *Phys.Lett.*, B288:342–347, 1992.
- [90] Luuk H. Karsten and Jan Smit. Lattice Fermions: Species Doubling, Chiral Invariance, and the Triangle Anomaly. *Nucl.Phys.*, B183:103, 1981.
- [91] A.D. Kennedy. Algorithms for dynamical fermions. 2006.
- [92] Daniel-Jens Kusterer, John Hedditch, Waseem Kamleh, Derek B. Leinweber, and Anthony G. Williams. Low lying eigenmodes of the Wilson-Dirac operator and correlations with topological objects. *Nucl.Phys.*, B628:253–269, 2002.
- [93] Jack Laiho and Ruth S. Van de Water. Pseudoscalar decay constants, light-quark masses, and B_K from mixed-action lattice QCD. 2011.
- [94] M.L. Laursen, J. Smit, and J.C. Vink. Small scale instantons, staggered fermions and the topological susceptibility. *Nucl.Phys.*, B343:522–540, 1990.
- [95] R. B. Lehoucq and D. C. Sorensen. Deflation techniques for an implicitly restarted arnoldi iteration. *SIAM J. Matrix Anal. Appl.*, 17:789–821, October 1996.
- [96] G. Peter Lepage and Paul B. Mackenzie. On the viability of lattice perturbation theory. *Phys.Rev.*, D48:2250–2264, 1993. *Phys. Lett. B*.
- [97] M. Luscher and P. Weisz. Computation of the Action for On-Shell Improved Lattice Gauge Theories at Weak Coupling. *Phys.Lett.*, B158:250, 1985.

- [98] M. Luscher and P. Weisz. On-Shell Improved Lattice Gauge Theories. *Commun.Math.Phys.*, 97:59, 1985.
- [99] Martin Luscher, Rajamani Narayanan, Peter Weisz, and Ulli Wolff. The Schrodinger functional: A Renormalizable probe for nonAbelian gauge theories. *Nucl.Phys.*, B384:168–228, 1992.
- [100] Martin Luscher, Rainer Sommer, Peter Weisz, and Ulli Wolff. A Precise determination of the running coupling in the SU(3) Yang-Mills theory. *Nucl.Phys.*, B413:481–502, 1994.
- [101] Martin Luscher, Rainer Sommer, Ulli Wolff, and Peter Weisz. Computation of the running coupling in the SU(2) Yang-Mills theory. *Nucl.Phys.*, B389:247–264, 1993.
- [102] Martin Luscher, Peter Weisz, and Ulli Wolff. A Numerical method to compute the running coupling in asymptotically free theories. *Nucl.Phys.*, B359:221–243, 1991.
- [103] G. Martinelli, C. Pittori, Christopher T. Sachrajda, M. Testa, and A. Vladikas. A General method for nonperturbative renormalization of lattice operators. *Nucl.Phys.*, B445:81–108, 1995.
- [104] C. McNeile and Christopher Michael. Decay width of light quark hybrid meson from the lattice. *Phys.Rev.*, D73:074506, 2006.
- [105] I. Montvay and G. Munster. *Quantum fields on a lattice*. Cambridge University Press, 1994.
- [106] Silvia Necco and Rainer Sommer. The $N(f) = 0$ heavy quark potential from short to intermediate distances. *Nucl.Phys.*, B622:328–346, 2002.
- [107] H. Neff, N. Eicker, T. Lippert, John W. Negele, and K. Schilling. On the low fermionic eigenmode dominance in QCD on the lattice. *Phys.Rev.*, D64:114509, 2001.
- [108] Herbert Neuberger. A Practical implementation of the overlap Dirac operator. *Phys.Rev.Lett.*, 81:4060–4062, 1998.

- [109] Herbert Neuberger. Exactly massless quarks on the lattice. *Phys.Lett.*, B417:141–144, 1998.
- [110] Herbert Neuberger. More about exactly massless quarks on the lattice. *Phys.Lett.*, B427:353–355, 1998.
- [111] Holger Bech Nielsen and M. Ninomiya. No Go Theorem for Regularizing Chiral Fermions. *Phys.Lett.*, B105:219, 1981.
- [112] Konstantin Ottnad, Carsten Urbach, Chris Michael, and Siebren Reker. Eta and eta' meson masses from Nf=2+1+1 twisted mass lattice QCD. 2011.
- [113] B. N. Parlett. *The Symmetric Eigenvalue Problem*. Prentice-Hall, Englewood Cliffs, NJ, 1980.
- [114] Joseph Polchinski. Renormalization and Effective Lagrangians. *Nucl.Phys.*, B231:269–295, 1984.
- [115] H. David Politzer. RELIABLE PERTURBATIVE RESULTS FOR STRONG INTERACTIONS? *Phys. Rev. Lett.*, 30:1346–1349, 1973.
- [116] Richard J. Radke. A matlab implementation of the implicitly restarted arnoldi method for solving large-scale eigenvalue problems, 1996.
- [117] H.J. Rothe. Lattice gauge theories: An Introduction. *World Sci.Lect.Notes Phys.*, 74:1–605, 2005.
- [118] Thomas Schafer and Edward V. Shuryak. Instantons in QCD. *Rev.Mod.Phys.*, 70:323–426, 1998.
- [119] Klaus Schilling, Hartmut Neff, and Thomas Lippert. Computing the eta and eta' mesons in lattice QCD. *Lect. Notes Phys.*, 663:147–176, 2005.
- [120] Yigal Shamir. Chiral fermions from lattice boundaries. *Nucl.Phys.*, B406:90–106, 1993.
- [121] Stephen R. Sharpe. Future of Chiral Extrapolations with Domain Wall Fermions. 2007.

-
- [122] J. R. Shewchuck. An Introduction to the Conjugate Gradient Method without the Agonizing Pain. 1994.
- [123] Stefan Sint. One loop renormalization of the QCD Schrodinger functional. *Nucl.Phys.*, B451:416–444, 1995.
- [124] Stefan Sint. The Schrodinger functional in QCD. *Nucl.Phys.Proc.Suppl.*, 42:835–837, 1995.
- [125] R. Sommer. A New way to set the energy scale in lattice gauge theories and its applications to the static force and alpha-s in SU(2) Yang-Mills theory. *Nucl.Phys.*, B411:839–854, 1994.
- [126] D. C. Sorensen. Implicit application of polynomial filters in a k -step arnoldi method. *SIAM Journal on Matrix Analysis and Applications*, 13(1):357–385, 1992.
- [127] C. Sturm, Y. Aoki, N.H. Christ, T. Izubuchi, C.T.C. Sachrajda, et al. Renormalization of quark bilinear operators in a momentum-subtraction scheme with a nonexceptional subtraction point. *Phys.Rev.*, D80:014501, 2009.
- [128] Gerard 't Hooft. Symmetry Breaking Through Bell-Jackiw Anomalies. *Phys.Rev.Lett.*, 37:8–11, 1976.
- [129] Gerard 't Hooft and M.J.G. Veltman. Regularization and Renormalization of Gauge Fields. *Nucl.Phys.*, B44:189–213, 1972.
- [130] Tetsuya Takaishi. Heavy quark potential and effective actions on blocked configurations. *Phys.Rev.*, D54:1050–1053, 1996.
- [131] J. Viehoff et al. Improving stochastic estimator techniques for disconnected diagrams. *Nucl.Phys.Proc.Suppl.*, 63:269–271, 1998.
- [132] Jeroen C. Vink and Uwe-Jens Wiese. Gauge fixing on the lattice without ambiguity. *Phys.Lett.*, B289:122–126, 1992.
- [133] Pavlos M. Vranas. Chiral symmetry restoration in the Schwinger model with domain wall fermions. *Phys.Rev.*, D57:1415–1432, 1998.

- [134] W.E.Arnoldi. The principle of minimized iterations in the solution of the matrix eigenvalue problem. *Quarterly of Applied Mathematics*, 9:17–29, 1951.
- [135] D.H. Weingarten and D.N. Petcher. Monte Carlo Integration for Lattice Gauge Theories with Fermions. *Phys.Lett.*, B99:333, 1981.
- [136] Kenneth G. Wilson. Confinement of Quarks. *Phys.Rev.*, D10:2445–2459, 1974.
- [137] K.G. Wilson and John B. Kogut. The Renormalization group and the epsilon expansion. *Phys.Rept.*, 12:75–200, 1974.
- [138] Z.Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst. Templates for the Solution of Algebraic Eigenvalue Problems: a Practical Guide. 2000.
- [139] Yuriy Zhestkov. Domain wall fermion study of scaling in nonperturbative renormalization of quark bilinears and B(K). 2001.

Appendix A

Eigenvalues in the Unit Gauge

A.1 D_W and H_W

The Wilson Dirac operator given in chapter 1 equation 1.24 in the unit gauge is,

$$D_W(x, y) = (4 + m)\delta(x, y) - \frac{1}{2} \sum_{\mu=1}^4 (1 + \gamma_\mu)\delta(x - \hat{\mu}, y) + (1 - \gamma_\mu)\delta(x + \hat{\mu}, y) \quad (\text{A.1})$$

this can be Fourier transformed using,

$$\begin{aligned} \frac{1}{V} \sum_{x,y} e^{-ipx} \delta(x, y) e^{iqy} &= \frac{1}{V} \sum_x e^{-ix(p-q)} = \delta(p, q) \\ \frac{1}{V} \sum_{x,y} e^{-ipx} \delta(x \pm \hat{\mu}, y) e^{iqy} &= \delta(p, q) e^{\pm iq_\mu} \end{aligned} \quad (\text{A.2})$$

where V is the spatial volume.

$$\begin{aligned} D_W(p, q) &= \frac{1}{V} \sum_{x,y} e^{-ipx} D_W(x, y) e^{iqy} \\ &= (4 + m)\delta(p, q) - \frac{1}{2} \sum_{\mu=1}^4 (e^{-iq_\mu} + e^{iq_\mu}) + \gamma_\mu (e^{-iq_\mu} - e^{iq_\mu}) \delta(p, q) \\ &= \delta(p, q) \left[m + \sum_{\mu=1}^4 (1 - \cos(q_\mu)) + i \sum_{\mu=1}^4 \gamma_\mu \sin(q_\mu) \right] \\ &= \delta(p, q) \tilde{D}_W(q) \end{aligned} \quad (\text{A.3})$$

To fully diagonalize the momentum space Wilson operator write it in the chiral basis where,

$$\gamma_\mu = \begin{pmatrix} 0 & -i\sigma_\mu \\ i\sigma_\mu & 0 \end{pmatrix} \quad (\text{A.4})$$

$$\gamma_5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (\text{A.5})$$

the 2×2 Pauli matrices σ satisfy,

$$\sigma_\mu = \sigma_\mu^\dagger \quad (\text{A.6})$$

$$\{\sigma_\mu, \sigma_\nu\} = \delta_{\mu\nu}. \quad (\text{A.7})$$

Then with,

$$A = m + \sum_{\mu=1}^4 (1 - \cos(q_\mu)) \quad (\text{A.8})$$

$$B = \sum_{\mu=1}^4 \sigma_\mu \sin(q_\mu)$$

$$B^2 = \sum_{\mu=1}^4 \sin^2(q_\mu)$$

we have

$$\tilde{D}_W(q) = \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \quad (\text{A.9})$$

the eigenvalues of this are found from the roots of the characteristic polynomial and are,

$$\lambda_\pm = A \pm i\sqrt{B^2}, \quad (\text{A.10})$$

the eigenvalues are functions of q and the entire spectrum can be obtained by scanning q over all its allowed values. The matrix,

$$\tilde{H}_W = \gamma_5 \tilde{D}_W = \begin{pmatrix} A & B \\ B & -A \end{pmatrix} \quad (\text{A.11})$$

is Hermitian, see section 1.3, and its eigenvalues are,

$$\lambda_{\pm} = \pm\sqrt{A^2 + B^2}. \quad (\text{A.12})$$

We also need the eigenvectors of this matrix,

$$\tilde{H}_W v_{\pm} = \lambda_{\pm} v_{\pm} \quad (\text{A.13})$$

which are normalized to one for later convenience,

$$\begin{aligned} v_{\pm} &= \begin{pmatrix} x_{\pm} \\ y_{\pm} \end{pmatrix} \\ x_{\pm} &= \frac{A + \lambda_{\pm}}{N_{\pm}} \\ y_{\pm} &= \frac{B}{N_{\pm}} \\ N_{\pm} &= \sqrt{B^2 + (A + \lambda_{\pm})^2} \end{aligned} \quad (\text{A.14})$$

A.2 H_T

The Shamir kernel, after the same transformation into momentum space, becomes

$$\tilde{H}_T = \tilde{H}_W \frac{1}{2 + \gamma_5 \tilde{H}_W} = \begin{pmatrix} A & B \\ B & -A \end{pmatrix} \begin{pmatrix} 2 + A & B \\ -B & 2 + A \end{pmatrix}^{-1} \quad (\text{A.15})$$

which simplifies to,

$$\tilde{H}_T = \frac{1}{D} \begin{pmatrix} A(2 + A) + B^2 & 2B \\ 2B & -A(2 + A) - B^2 \end{pmatrix} = \begin{pmatrix} P & Q \\ Q & -P \end{pmatrix} \quad (\text{A.16})$$

$$\begin{aligned} D &= ((2 + A)^2 + B^2) \\ P &= \frac{1}{D}(A(2 + A) + B^2) \\ Q &= \frac{1}{D}(2B). \end{aligned} \quad (\text{A.17})$$

The eigenvalues are,

$$\lambda_{\pm} = \pm \sqrt{P^2 + Q^2} \quad (\text{A.18})$$

and the eigenvectors as before, but with $A \rightarrow P$ and $B \rightarrow Q$

A.3 D_{ov} and H_{ov}

We derived the operator $D_{ov} = \gamma_5 H_{ov}$ in section 1.6. The massless D_{ov} is given by,

$$D_{ov} = \frac{1}{2} [1 + \gamma_5 f(H)] \quad (\text{A.19})$$

where,

$$\begin{aligned} f(T) &= \tanh \left(-\frac{L_s}{2} \log |T| \right) \\ T &= \frac{1 - H}{1 + H}. \end{aligned} \quad (\text{A.20})$$

and the kernel H is H_T or H_W . The same Fourier transformation as before takes D_{ov} into an almost diagonal form, after which we need to evaluate a function of a Hermitian matrix. Diagonalize H in the basis of its eigenvectors,

$$\begin{aligned} V &= (v_+ v_-) \\ V^\dagger H V &= \Lambda = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix} \end{aligned} \quad (\text{A.21})$$

then the spectral theorem gives,

$$f(H) = V f(\Lambda) V^\dagger = \begin{pmatrix} a & b \\ b & d \end{pmatrix} \quad (\text{A.22})$$

note this is Hermitian. The matrix $f(H)$ can be calculated from the information given in the previous sections. We have to specify a value for the domain wall height which we do by setting the mass term in H to $m = -M_5 = -1$. The eigenvalues of D_{ov} are,

$$\mu_{\pm} = \frac{1}{2} \left[1 + \frac{1}{2} \left((a - d) \pm \sqrt{(a + d)^2 - 4b^2} \right) \right] \quad (\text{A.23})$$

and the eigenvalues of H_{ov} are

$$\mu_{\pm} = \frac{1}{4} \left[(a + d) \pm \sqrt{(2 + a - d)^2 + 4b^2} \right]. \quad (\text{A.24})$$

We can also obtain the eigenvalues of the exact overlap by substituting

$$f(H) \rightarrow \text{sign}(H) = \frac{H}{\sqrt{H^2}} \quad (\text{A.25})$$

then

$$\begin{aligned} \text{sign}(H) &= V \text{sign}(\Lambda) V^{\dagger} = v_+ v_+^{\dagger} - v_- v_-^{\dagger} \\ &= \begin{pmatrix} x_+ x_+ - x_- x_- & x_+ y_+ - x_- y_- \\ x_+ y_+ - x_- y_- & y_+ y_+ - y_- y_- \end{pmatrix} = \begin{pmatrix} a & b \\ b^{\dagger} & d \end{pmatrix}. \end{aligned} \quad (\text{A.26})$$

From equation A.14 we then have,

$$(a + d) = (x_+ x_+ + y_+ y_+) - (x_- x_- + y_- y_-) = 1 - 1 = 0 \quad (\text{A.27})$$

letting $\lambda = \sqrt{A^2 + B^2}$ so $\lambda_{\pm} = \pm \lambda$ we have

$$\begin{aligned} b &= x_+ y_+ - x_- y_- \\ &= AB \left(\frac{1}{N_+^2} - \frac{1}{N_-^2} \right) + \lambda B \left(\frac{1}{N_+^2} + \frac{1}{N_-^2} \right) \\ &= \frac{4B\lambda}{N_+^2 N_-^2} (\lambda^2 - A^2) \end{aligned} \quad (\text{A.28})$$

and

$$\begin{aligned} (a - d) &= (x_+ x_+ - y_+ y_+) - (x_- x_- - y_- y_-) \\ &= (\lambda^2 + A^2 - B^2) \left(\frac{1}{N_+^2} - \frac{1}{N_-^2} \right) + \lambda 2A \left(\frac{1}{N_+^2} + \frac{1}{N_-^2} \right) \\ &= \frac{8A\lambda}{N_+^2 N_-^2} (\lambda^2 - A^2). \end{aligned} \quad (\text{A.29})$$

hence the eigenvalues of D_{ov} are,

$$\frac{1}{2} \left[1 + \frac{4\lambda}{N_+^2 N_-^2} (\lambda^2 - A^2) \left(A \pm i\sqrt{B^2} \right) \right] \quad (\text{A.30})$$

substituting in,

$$N_{\pm}^2 = 2\lambda(\lambda \pm A) \tag{A.31}$$

we have the eigenvalues of D_{ov} in the unit gauge,

$$\frac{1}{2} \left[1 + \frac{A \pm i\sqrt{B^2}}{\sqrt{A^2 + B^2}} \right] \tag{A.32}$$

which shows how the overlap construction projects the eigenvalues of D_W onto the unit circle.