Study of Instanton Physics in Lattice QCD

by

Taras Ivanenko

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

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1997

© Massachusetts Institute of Technology 1997. All rights reserved.

Author .....
Department of Physics
May 15, 1997

Certified by .....
John William Negele
William A. Coolidge Professor of Physics
Thesis Supervisor

Accepted by .....
George Koster
Chairman, Departmental Committee on Graduate Students

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

JUN 09 1997 Science

LIBRARIES

Study of Instanton Physics in Lattice QCD

by

Taras Ivanenko

Submitted to the Department of Physics
on May 15, 1997, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

One of the most interesting parts of QCD dynamics is the non-perturbative region of long-range excitations and this thesis explores several semi-classical aspects of this region associated with instantons. The instanton distribution in lattice QCD has been calculated with and without internal fermions and no statistically significant difference between the two distributions was found for a sea quark mass $m_q \approx 100$MeV. Since the instanton distribution is observed by using a relaxation process, the effects of lattice artifacts and pair annihilation during relaxation are investigated.

The role of the low-eigenvalue modes in the Hybrid Monte-Carlo dynamics has been studied and it is shown that the efficiency of this method can be significantly increased by using a few recent vectors to approximate the solution of the inverse Dirac operator. The algorithm to find the low-eigenvalue modes of the Dirac operator was developed and used to study the spectrum of the Dirac–Wilson operator on the lattice. Restricting the Dirac operator to the complement of the few lowest modes is shown to completely eliminate critical slowing down in the Conjugate Gradient method.

Point-to-point correlation functions were calculated using the full propagator and the propagator in the subspace of low-eigenvalue modes of the Dirac operator. The "truncated" correlation function reproduces the behaviour of the full correlation function at large distance. This result agrees with the prediction of the instanton liquid model that quark dynamics in QCD can be described by instanton zero modes.

Thesis Supervisor: John William Negele
Title: William A. Coolidge Professor of Physics

Acknowledgments

My study at the Center for Theoretical Physics at MIT was a pleasant and enriching experience and I want to thank all people who made it so. I am grateful to my advisor John Negele for bringing me to MIT and supporting my research. My first advisors, Mikhail Polikarpov, Leonid Kondratyuk and Yuri Simonov introduced me to the Quantum Field Theory and guided me during my years of study in Moscow at MPhTI and ITEP. I would like to thank Jan Smit and Pierre van Baal for helping me to survive in Amsterdam on my way to MIT. I am obliged to Richard Brower and Edward Shuryak for many interesting discussions we had.

The research described in this thesis was partially supported by U.S. Department of Energy (D.O.E.) under cooperative research agreement DE-FC02-94ER40818. The calculations were done initially on CM5 at MIT within Project Scout and continued on SUN Enterprise 5000 Xolas facility.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Quantum Chromo–Dynamics theory of strong interactions is the conventional way of describing strongly interacting matter: quarks, gluons, mesons and baryons. With the success of the theory in the high-energy region came the challenge of describing the intermediate– and low– energy physics from the first principles of QCD. Due to the asymptotic freedom of QCD, the high-energy processes can be successfully described within perturbation theory. The same asymptotic freedom leads to the breakdown of perturbation theory in the low-energy region where other methods have to be developed to explain the hadron structure, quark confinement and the rest of the experimental data.

Instanton models [73] provide the semiclassical description of QCD dynamics. Although instanton models have difficulties describing confinement, they give a reasonably good description of chiral symmetry breaking and hadron structure. The range of validity of the semi-classical instanton approximation is complementary to the perturbation theory range. While perturbation theory produces an expansion in the coupling constant g^2, the relevant parameter for instantons is the one-instanton action $S_I = \frac{8\pi^2}{g^2}$. The instanton liquid model takes as an input the phenomenological parameters, instanton distribution or some other quantities, which can be determined from either experiments or lattice calculations.

The lattice regularization method is one of the possible methods of perturbative regularization but it also provides a framework for *ab initio* calculation of all QCD quantities. In theory, the lattice results are equal to those obtained by other regularization methods in continuum limit. In practice, limited computer resources restrict the size of the lattice and the continuum limit is defined only by extrapolating results from finite lattice spacing and finite lattice volume with statistical errors arising from the Monte-Carlo integration methods. At present, lattice calculations provide the only method of studying non-perturbative effects in QCD from first principles.

This work addresses the long-range excitations in lattice QCD. The comparison of the lattice results with the results of instanton models establishes the link between those models and full QCD.

The study of instantons on the lattice can also provide insight on the dynamics of lattice QCD algorithms.

One of the key tool in studying QCD is point-to-point correlation functions for quark currents. Many correlation functions can be determined experimentally and any successful model has to describe them. The instanton gas or liquid models describe the correlation functions by approximating the propagator with the propagator in the subspace of the instanton zero modes. This approximation provides a good description of correlation functions and the same approximation was studied on the lattice, using the exact computed spectrum of the Dirac operator on the lattice.

The algorithms developed for studying instanton physics can also be applied to other problems of lattice QCD and beyond lattice gauge theories.

Chapter 2 reviews the lattice regularization method (Sec. 2.1), defines the lattice gauge fields (Sec. 2.2) and fermions (Sec. 2.3), specifies the action (Sec. 2.4) and defines the Hybrid Monte-Carlo algorithm (HMC) used for generating the lattice configurations (Sec. 2.5).

Chapter 3 describes the method of accelerating the computations in HMC method by using the analytical properties of the HMC evolution. Taking into account the slow evolution of low-eigenvalue modes during HMC trajectory we have achieved significant increase in the efficiency of the algorithm.

Chapter 4 addresses the instanton distribution in lattice QCD. It begins with a review of instanton properties and instanton models for QCD dynamics (Sec. 4.1). Section 4.2 describes the relaxation algorithm used to suppress the quantum noise in the dynamically generated configurations and Section 4.3 describes the procedures used to identify the lattice instantons. In Section 4.4 the effects of the relaxation process on the isolated instantons are studied to determine the distortions of the instanton distribution introduced by relaxation. Section 4.5 reports the results of the measurements of instanton distribution in QCD with and without internal fermions and discuss the results of other groups on this subject.

Chapter 5 presents the spectrum of the Dirac–Wilson operator on the lattice. Section 5.1 discusses the general properties of the lattice Dirac-Wilson operator and, in Section 5.2, the spectrum for free fermions is studied. Section 5.3 investigates in detail the low-eigenvalue part of the spectrum in the field of a single isolated instanton. The results of those sections establish the framework to study the effects of the instantons in the spectrum of Dirac-Wilson operator in lattice QCD. Section 5.4 discusses the effects of the instanton interaction on the spectrum and the role of the Wilson term in the Dirac operator. Section 5.5 presents the low-eigenvalue part of the spectrum in the dynamically generated field configuration and compares the lattice computation with continuum results. In Section 5.6, the computed eigenvectors are used to study the role of low-eigenvalue part of the spectrum on the convergence of Conjugate Gradient algorithm. Separating the low-eigenvalue part of the spectrum in the Dirac operator completely eliminates the critical slowing down. To study the effects of the instantons on quark propagation, the quark current correlation function was

computed by restricting the quark propagator spectral representation to few modes (Sec. 5.7). The results show that essential part of the quarks dynamics is contained in the low-eigenvalue part of the spectrum.

Chapter 6 gives the summary of main results described in this thesis.

Appendix A defines the gamma-matrices representation used in this work and summarizes the properties of t'Hooft's η-symbols. Appendix B describes the Even-Odd pre-conditioning method and the k-step Arnoldi method used for computing the eigenvectors.

Chapter 2

Definitions and Algorithms

2.1 Lattice Regularization

Regularization is essential to a definition of any quantum field theory. Lattice regularization is one of the possible methods to render the theory finite and start the renormalization procedure. First, some mesh is defined in the 4-dimensional Euclidean space. Any mesh breaks the Poincarè group invariance to some discrete group and the full invariance group is recovered only in the continuum limit. In the case of a regular rectangular mesh, this residual group is a product of 4-dimensional cubic group and finite translations. Next, the continuum fields are replaced by lattice variables and continuum operators are replaced by lattice operators acting on lattice variables. The lattice variables and operators are defined in the following Sections.

The continuum limit of the lattice field theory is reached when the correlation length ξ, expressed in lattice units a, diverges. At this point lattice artefacts ($\sim a/\xi$) disappear and the lattice theory can be related to the continuum theory. In lattice QCD this point is at $\beta = \frac{2N_c}{g^2} \to \infty$ and at this point all lattice variables, expressed through lattice spacing a, follow the perturbative asymptotic scaling law

$$a\Lambda_{\rm L} = \exp\left(-\frac{1}{2b_0 g_0^2}\right)\left(b_0 g_0^2\right)^{-b_1/2b_0^2}, \tag{2.1}$$

where

$$b_0 = \frac{1}{16\pi^2}\left(11N_c/3 - 2N_f/3\right) \text{ and } b_1 = \frac{1}{(16\pi^2)^2}\left(34N_c^2/3 - 13N_cN_f/3 + N_f/N_c\right) \tag{2.2}$$

To make a correspondence with the continuum theory we need some observable which can be measured both on the lattice and in experiments (or deduced from the experimental data). The string tension $\sigma = (440\text{MeV})^2$ and ρ-meson mass $m_\rho = 770\text{MeV}$ are possible choices of variables used to set the scale of the lattice spacing a and quark mass q.

Additional complications arise from the lattice boundary effects. Since the lattice in numerical computations can not be infinite or even very large, the scaling window $a \ll R \ll L$ is usually not very big. The boundaries also introduce additional violations of rotational invariance.

2.2 Gauge Fields

In the standard way of representing the gauge fields on the lattice [18, 19] each link of the lattice carries the element of the gauge group $U_{x,\mu} \in SU(N)$. The index x denotes the site on the lattice and μ is the direction of the link. In this work all links point in positive directions, that is, connect points x and $x + \hat{\mu}$ where $\hat{\mu}$ is a unit lattice vector in direction μ. In the continuum limit the link field $U_{x,\mu}$ is a path-ordered integral of the gauge field $A_\mu(z)$ along the straight line connecting x and $x + \hat{\mu}$.

$$U_{x,\mu} = P \exp\left(i \int_{x}^{x+\hat{\mu}} A_\mu(z) dz \right) \approx 1 + iaA_{x,\mu} \tag{2.3}$$

The trace of a product of link matrices $U_{x,\mu}$ along the closed path is gauge-invariant and makes a building block for constructing gauge field-related observables.

2.3 Fermions

The fermion fields on the lattice are naturally represented by their values on the sites of the lattice ψ_x belonging to some appropriate Grassman algebra.

The simplest, so called "naive", discretization of Dirac operator is obtained by replacing the derivative by the central difference. This operator has a lattice correction proportional to a^2 while the one-sided difference operator has corrections $\sim a$.

$$D\psi_x = m\psi_x + \sum_\mu \frac{1}{2}\gamma_\mu(\psi_{x+\hat{\mu}} - \psi_{x-\hat{\mu}}) \tag{2.4}$$

The spectrum of the "naive" Dirac operator for free fermions can be obtained by simple Fourier transformation

$$\lambda_n = m + i\sqrt{\sum_\mu \sin(\frac{2\pi n_\mu}{L_\mu})^2}, \tag{2.5}$$

where $n_\mu = -L_\mu/2 \ldots L_\mu/2$. The physical part of a spectrum is the region where $\lambda \to 0$ when $L \to \infty$ and the continuum dispersion relation $\lambda = m + i\|p\|$ is recovered. It is easy to see that naive fermions have not one but 16 such regions, one in the vicinity of $n_\mu = 0$ and $n_\mu = L_\mu/2$ for each μ, corresponding to 16 species of massless fermions in continuum limit. Those extra fermions are called "doublers". Only one of the doubler modes is smooth when $a \to 0$, the others have a sawtooth structure $\sim (-1)^x$.

Note that the word "spectrum" describes the spectrum of the Dirac operator, not the energy spectrum $E(p)$. The two notions of spectrum are related but I use only the operator spectrum throughout this thesis. The spectrum of a Dirac operator can be represented on a complex plane where real and imaginary axes do not have separate physical meaning.

One solution to the doubler problem, proposed by Wilson [84], is to add a term to the action which gives extra mass to the doublers while not affecting the physical region $n_\mu \approx 0$. The simplest such term is a Laplacian $r\Delta$. The full Wilson-Dirac operator is conventionally written as

$$D\psi_x = \psi_x - \kappa \sum_\mu \left[(r - \gamma_\mu) U_{x,\mu} \psi_{x+\hat{\mu}} + (r + \gamma_\mu) U^\dagger_{x-\hat{\mu},\mu} \psi_{x-\hat{\mu}} \right] \tag{2.6}$$

The coefficient r is arbitrary and usually set to 1 for simplicity. The hopping constant κ is related to a fermion mass by

$$\frac{1}{\kappa} - \frac{1}{\kappa_c} = 2m, \tag{2.7}$$

where κ_c is a critical κ, and $\kappa_c = 1/8$ for free fermions. In the case of interacting fields this equation is used to define the quark masses after κ_c is determined separately (see the example of determining κ_c in Chapter 5.

The spectrum of Wilson–Dirac operator (Eq. 2.6) even for free fermions

$$\lambda_n = 1 - 2\kappa \left(\sum_\mu \cos(\frac{2\pi n_\mu}{L_\mu}) \right) + 2i\kappa \sqrt{\sum_\mu \sin(\frac{2\pi n_\mu}{L_\mu})^2} \tag{2.8}$$

has a correct continuum limit but is complex in general. Since the Wilson term is not chirally invariant, the chiral symmetry of QCD is recovered only in the continuum limit $\kappa \to \kappa_c$. I will discuss this operator in more details in Section 5.1.

2.4 Lattice Action

The action for the gauge fields is a functional of the gauge-invariant closed loops. The simplest choice is the Wilson action

$$S_W[U] = \sum_{x,\mu\neq\nu} \left(1 - \frac{1}{N_c} \mathrm{Tr} P_{x,\mu,\nu} \right), \tag{2.9}$$

where the plaquette $P_{x,\mu\nu}$ is a product of 4 links along the side of elementary square. It is possible to use more terms to the action adding bigger loops when necessary [77, 78], adjusting the coefficients of different terms to reduce lattice artefacts. Both the Wilson action and an improved action have been used in this work (see Section 4.2).

The actions for the gauge fields and fermions define the partition function of the theory

$$\mathcal{Z} = \int DUD\bar{\psi}D\psi \exp\left(-\beta S[U] + \bar{\psi}(D+m)\psi\right) = \int DU \mathrm{e}^{-\beta S[U]} \left[\det(D+m)\right]^{N_f} \tag{2.10}$$

Calculating the partition function by direct integration is hopeless because of the enormous number of degrees of freedom and but it can be evaluated by Monte-Carlo integration. The integration over the phase space with a weight (Eq. 2.10) is replaced by the average over the set of fields $\{U^{(i)}\}$ generated by a Markov process with correct probability $P[U] = e^{-S[U]}$. This replacement is valid if the process is ergodic and satisfies the detailed balance principle [18, 19]:

$$P[U]P(U \to U') = P[U']P(U' \to U) \tag{2.11}$$

The action for the pure gauge theory is well localized and the statistical sample of configurations can be generated by local updates. Two examples of such algorithms are the Cabibbo-Marinari (heat-bath) algorithm and the Metropolis algorithm [18, 19]. By using only the pure gluon action and omitting the determinant, the effects of internal fermion loops are ignored. This approximation is called the quenched or valence approximation. I also use the term "full" to denote the Monte–Carlo process and configurations with internal fermions.

2.5 Hybrid Monte-Carlo Algorithm

Since fermions are represented by anti-commuting variables, there is no practical way to define them on a computer. After the fermions are integrated out of the partition function (Eq. 2.10), the resulting determinant introduces highly non-local interactions of the gauge fields. In the quenched (or valence) approximation, all of the effects of the determinant interaction are ignored. In physical terms, this approximation corresponds to neglecting the internal fermion loops in all diagrams. At present, typical methods for going beyond the quenched approximation for lattice QCD involve an increase of several orders of magnitude in computational time — thus severely limiting statistics on large lattices. Consequently, finding a more efficient approach to the inclusion of internal fermion loops poses a major challenge for the next generation of lattice simulations [53].

The most successful approach to date for generating full QCD configurations is the so-called Hybrid Monte Carlo (HMC) algorithm [6, 15, 27, 28, 29, 30, 37, 39, 40, 41, 54, 56]. This algorithm has the advantage that, apart from numerical round off, it provides an exact Markov process for generating QCD configurations. Starting from the Euclidean partition function for QCD (Eq. 2.10) with even number of flavours of Wilson Fermions and using the identity

$$\left[\det(M)\right]^{N_f} = \left[\det(M^\dagger M)\right]^{N_f/2} \tag{2.12}$$

we may rewrite the determinant as Gaussian integral over a set of bosonic "pseudo-fermion" fields φ_x. The HMC algorithm also requires additional canonical (angular) momentum coordinates, $\mathrm{E}_\mu(x)$, conjugate to the gauge fields $A_\mu(x)$ on each link $(x, x+\hat{\mu})$. Combining these steps the QCD partition function is now rewritten as

$$Z = \int DUDED\varphi e^{-H(U,E,\varphi)} \tag{2.13}$$

with an entirely bosonic action given by

$$H(U,E,\varphi) = \tfrac{1}{2}\mathrm{TrE}^2 + S_g(U) + \varphi^\dagger [M^\dagger M]^{-1}\varphi \tag{2.14}$$

Now the QCD partition function looks like the partition function of a Hamiltonian system.

The standard HMC algorithm alternates Gaussian updates with Hamiltonian evolution in a "fifth time" coordinate t to achieve detailed balance and ergodicity, the evolution part of the algorithm is called a Molecular Dynamics algorithm (MD) for historical reasons. At the beginning of each Hamiltonian trajectory the momenta, E, and "pseudo-fermion" fields, $b = M^{\dagger^{-1}}\varphi$, are chosen as independent Gaussian random variables. Next the gauge fields are evolved for a MD time T using the Hamiltonian equations of motion with fixed values of the "pseudo-fermion" fields φ. Since the updates at the beginning of each trajectory do not depend on the history of the evolution, the algorithm realizes the Markov process for the gauge fields U. By the properties of the Hamiltonian evolution the probability distribution,

$$\mathcal{P}(U) = \int D\psi D\bar{\psi} e^{-S_g(U)+\bar{\psi}M(U)\psi} , \tag{2.15}$$

is invariant during the evolution.

In practice, the Hamiltonian evolution must be approximated by an integration scheme with a finite step size δt. Consequently, at the end of each MD trajectory, a Metropolis accept-reject test is introduced to remove integration errors which would otherwise corrupt the action being simulated. In addition, it is also essential that the MD integration method exactly obey time-reversal invariance to avoid violating detailed balance. Clearly small violations of detailed balance are inevitable due to round off errors. Current practice usually employs the leapfrog integration scheme in single precision arithmetic. This scheme introduces an error in the action proportional to δt^2. Better integration schemes could allow to increase δt and, therefore, reduce auto-correlation time. Also, with standard even-odd pre-conditioning [26, 39, 40, 41], the full Dirac operator D is replaced by an even–to–even operator M_{ee} (see Sec. B.1). In this case, only pseudo-fermions φ on the even sites have to be introduced and the matrix,

$$A = M_{ee}^\dagger M_{ee}, \tag{2.16}$$

is used in the algorithm.

There are many small variations on this basic method, but common to all HMC algorithms is the need to accurately integrate the equations of motion, calculating the force on U due to the pseudo-fermions at each time step t_n. The calculation of force on the gauge field requires solving the Dirac equation,

$$A(t_n)\chi(t_n) = \varphi \tag{2.17}$$

over and over again for the propagator $\chi(t)$, where $A(t) \equiv M(U)^\dagger M(U)$. Technically, this is achieved by starting with a trial value χ_{trial} and iteratively solving Equation (2.17) for $\chi(t)$. The solution of Equation 2.17), which is usually obtained using the conjugate gradient (CG) method, is the most computationally expensive part of Hybrid Monte Carlo algorithm.

Chapter 3

Accelerating HMC with "Chronological" Inverter

At each time step of the Hybrid Monte-Carlo method the Dirac matrix must be inverted to calculate the force due to fermion loops. This inversion is the most time-consuming part of the full QCD algorithm. The Dirac operator is represented as a $12V$ by $12V$ sparse square matrix, where the number of space-time lattice sites (or volume V) is on the order of 10^{5-6}. With the lattice volumes used in recent simulations, the Dirac matrix may account for 90% or more of the computation time. Moreover, due to the Dirac inverter, the problem scales as $(1/a)^{6-7}$. Therefore any acceleration of the Dirac inversion will result in nearly a proportional net gain in the performance of the full QCD simulations.

There are several ways to increase the efficiency of the algorithm. Better inversion algorithms are always an option worth exploration [33, 45, 55]. Since the detailed balance of the HMC process is maintained by the final Metropolis accept/reject step, the MD trajectory need not be an exact approximation to the Hamiltonian evolution. In this method the exact (up to machine precision) inversion of the Dirac operator is replaced by some approximation to it still maintaining the reversibility [32, 45]. At the end, the time saved in reducing the number of inversions is traded for the lower acceptance rates and/or longer auto-correlation time.

On the other hand, if the inverse is solved exactly (in machine precision), the detailed balance is preserved, regardless of the details of the algorithm. The conjugate gradient method always starts with some initial vector, the first guess to a solution. Being able to guess better, the total time to reach convergence could be reduced. In our approach, the trial vector is "extrapolated" from a set of the previous solutions of the Dirac inversion. By virtue of the system's smooth evolution in MD time, the solution $\chi(t)$ must also change smoothly and it should be possible to use information from the recent past to perform the next inverse more efficiently [12, 13]. We call the method of guessing

the solution from the set of the past solutions the "chronological" inverter.

Consider the dynamics of the eigenmodes of the Dirac operator. Since the convergence of the CG algorithm is determined by the condition number $|\lambda_{max}/\lambda_{min}|$ and the lowest eigenmodes are the last to converge [36], a good guess in the subspace of the lowest modes can significantly push CG to convergence. The lowest modes are also the least affected by small changes in the operator. On the other hand, those modes are enhanced by the inverse and therefore the subspace of the few past solutions approximates the subspace of the lowest eigenmodes of the Dirac operator. See also Section 5.6 for further discussion of the relations between eigenmodes and the conjugate gradient algorithm.

Several possibilities can be taken into account. The first obvious step, employed frequently in HMC codes, is simply to use the previous solution as the starting trial solution. Some groups have even used a linear extrapolation of the last two solutions [37, 39, 40, 41]. The next natural step is to use a high order polynomial extrapolation. It was shown that the polynomial extrapolation gives a robust speed up in the CG [13]. Motivated by this success, we fixed the trial vector as a linear combination of the N previous solutions by minimizing the residual in the norm of the matrix (see Sec. 3.3). This approach, which we call the Minimal Residual Extrapolation (MRE) method, further improved the performance.

The generic problem of a chronological sequence of matrix inversions is not unique to HMC for QCD. For example, in many fluid dynamics codes, one must solve Poisson's equation for the pressure field as an inner loop for integrating the Navier–Stokes equation [44]. We believe the ideas developed here will be useful in many of these related applications.

3.1 Details of the simulations

The work on Chronological Inverter and Hybrid Monte–Carlo code was done in a BU–MIT collaboration including Richard Brower (Boston U,MIT), Andreas Ruben Levi (Boston U) and Kostas Orginos (Brown U) on the CM-5 at MIT. We had available two CM-5, one with 128 nodes and the other with 64 nodes. We used the configurations generated for the MIT–BU lattice collaboration. We tested various values of the molecular dynamics step δt for 5 or more independent thermalized full QCD configurations, separated by approximately 100 MD trajectories.

In our tests we performed extensive simulations on a 16^4 lattice at $\beta = 5.5$ and $\kappa = 0.160$. These values of parameters give the lattice spacing $a_\rho = 0.11$fm, the quark mass $m_q \approx 100$MeV and the pion mass $m_\pi \approx 630$MeV (see Section 5.7 and [41, 39]). We choose δt in a typical window for actual simulations ($0.002 \leq \delta t \leq 0.015$). Some simulations were performed with a lighter quark mass and we obtained similar results [13], although the magnitude of the improvement depended slightly on the lattice parameters.

We used a standard CG method, and we iterated until the normalized squared residual,

$$R = \frac{|M_{ee}^{\dagger} M_{ee} \chi - \varphi|^2}{|\chi|^2}, \tag{3.1}$$

reached a given value. We choose the stopping condition, in much the same spirit as references [39, 40, 41], so that the error in computing the ΔS, used in the Metropolis accept-reject step, should on average be less than 1%. However it is not really known how this error propagates to physical quantities.

A critical issue is the requirement to converge accurately to the final solution for χ. Unless you start from a value of χ which is *independent* of past values (e.g. $\chi = \varphi$), the initial guess introduces an element of non-reversible dynamics and therefore violates detailed balance. Even in one of the most commonly used HMC methods, starting from the last value of χ (or a constant extrapolation), failure to converge to the correct inverse introduces a bias that breaks time reversal invariance. In principle, only if the CG has converged exactly to the fixed point of the conjugate gradient iterations, will there be no violation of time reversal invariance [46, 31]. In practice, we have measured this source of error by explicitly reversing the MD dynamics for various values of the stopping conditions (see Eq. 3.1). In Figures 3-1, 3-2 we show examples for $R = 10^{-10}$, $R = 10^{-11}$, $R = 10^{-12}$, $R = 10^{-13}$, $R = 10^{-14}$ and $R = 10^{-15}$. In Figure 3-1 we can see the action variation (total, fermion, gauge and momenta) in a trajectory that goes forward in MD time 100 steps and then is reversed for 100 steps. We have subtracted the initial action, so if the dynamics conserves energy, the total action should remain zero and if the dynamics is reversible, the total action should be symmetric around the MD time $t = 100$.

Our data demonstrates that simulations with $R \geq 10^{-12}$ are not sufficient to ensure a reversible dynamics. In Figure 3-2, we plot the total action difference between symmetric points in a forward-backward trajectory. If the dynamics is reversible, this difference should be zero. In this figure we can see that for $R > 10^{-13}$ the dynamics is not reversible. For $R = 10^{-13}$ the action difference observed is of order of 10^{-2} where the total action is of order 10^7, which means that we are at the boundaries of single precision accuracy. However, for $R = 10^{-14}$ we are well bellow the single precision accuracy. Smaller action variations can not be detected. Consequently, we chose to run our simulation using $R = 10^{-14}$ as the conjugate gradient stopping criterion. Running the conjugate gradient beyond $R = 10^{-14}$ eventually leads to divergence, although a few restarts quickly reduce the residue by an additional order of magnitude.

We measured the total number of CG steps needed to invert the Dirac Matrix to a given precision. The CPU time for a single leapfrog trajectory is practically proportional to the number of CG steps. However, the true performance of the method should be judged by the total number of CG steps needed to evolve the system for a fixed MD time, i.e. the total number of CG steps needed to

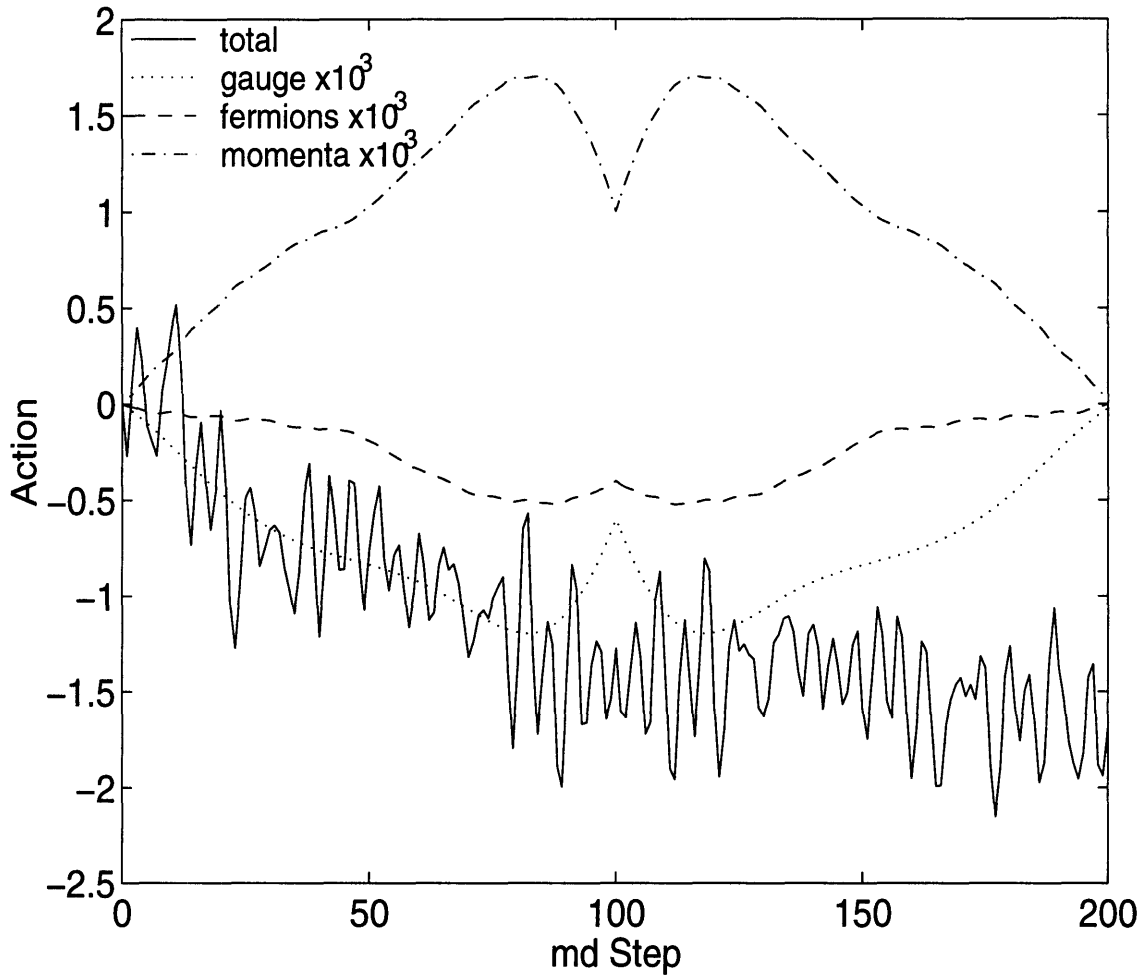

Figure 3-1: Action variation in a forward-backward trajectory. The initial action has been subtracted. $R < 10^{-10}$ was used as a stopping criterion and $\delta t = 0.010$.

compute a trajectory of time length T. We define the quantity,

$$CT = \frac{N_{CG}}{\delta t}, \tag{3.2}$$

which is proportional to the "computational time". Moreover, since a smaller step size will improve the acceptance rate, the trade-off may be even more favorable for small δt, and this effect should also be included in an overall measure of efficiency. However, because we always worked with δt in a region of very good acceptance rates ($> 99\%$), we have not included this effect in our estimate of overall efficiency in our simulations.

3.2 Chronological Inverter for QCD

We began our investigation by considering a trial solution, $\chi(0)$, of the new Dirac matrix $A(0)$, as a linear superposition of old solutions,

$$\chi_{trial} = c_1\chi(t_1) + c_2\chi(t_2) + \cdots + c_N\chi(t_N). \tag{3.3}$$

To simplify our notation, we will always suppose that the new inverse is computed at $t = 0$, given the past values at t_i, with $\cdots < t_3 < t_2 < t_1 < 0$. In practice, this is usually a regular series of values $t_n = -n\ \delta t$ with an integration step δt.

To date HMC simulations have used either the past solution as a trial solution, $\chi_{trial} = \chi(t_1)$, or a linear [37, 39, 40, 41] extrapolation, $\chi_{trial} = 2\chi(t_1) - \chi(t_2)$. It is natural to try to improve the estimate for the trial value by using higher order polynomial extrapolations. For example, if one uses an N-th order polynomial to fit $N + 1$ past values, the coefficients are given by

$$c_k = (-1)^{k-1}\frac{N!}{k!(N-i)!}. \tag{3.4}$$

Although substantial improvements can be made using higher order polynomial extrapolations, the method breaks down for large N ($N > 5$) [12, 13].

Since the conjugate gradient method is in fact just a minimal residual technique confined to the Krylov subspace spanned by vectors $A^{j-1}\chi_{trial}$ [36], why not start by examining a "smarter" subspace based on past success for nearby times? In this spirit, we suggest determining the coefficients c_n by minimizing the functional,

$$\Psi[\chi] = \chi^\dagger M^\dagger M\chi - \varphi^\dagger\chi - \chi^\dagger\varphi, \tag{3.5}$$

which is the same functional minimized by the Conjugated Gradient method itself. This corresponds to the minimization of the norm of the residual in the norm of the inverse matrix,

$$r^\dagger\frac{1}{M^\dagger M}r = \chi^\dagger M^\dagger M\chi - \varphi^\dagger\chi - \chi^\dagger\varphi + b^\dagger b, \tag{3.6}$$

in the subspace spanned by $\chi_i \equiv \chi(t_i)$, where $r = \phi - M^\dagger M\chi$ and $b \equiv (M^\dagger)^{-1}\varphi$. Therefore we call this method "Chronological Inverter by Minimal Residual Extrapolation" or MRE. The minimization condition reduces to

$$\sum_{j=1}^{N}\left(\chi_i^\dagger M^\dagger M\chi_j\right)c_j = \chi_i^\dagger\varphi. \tag{3.7}$$

The only technical problem is that this system can be poorly conditioned because the past solutions χ_i differ from each other by order δt. However, if properly handled, this instability should not

affect the resulting minima of the quadratic form in the span of the vectors, χ_i. We have explored other methods of extrapolations to some extent, in particular other slightly different definitions of residue (not in the norm of the inverse matrix) were examined, but nothing appears at present to out-perform the Minimal Residual Extrapolation method.

3.3 Chronological Inverter by Minimal Residual Extrapolation

We implemented this method by calling a routine that minimizes the residual in the subspace of the past solutions, right before each CG in the usual leapfrog. This routine has as inputs the set of past vectors χ_i, $i = 1 \ldots N$ and returns a vector χ_{trial} that minimizes $\Psi[\chi]$ in the $span(\chi_i)$:

MRE Algorithm

- Construct an orthogonal basis v_i in the $span(\chi_i)$ using the Gram-Schmidt procedure.
- Form the sub-matrix $G_{nm} = v_n^\dagger M^\dagger M v_m$, and the vector $b_n = v_n^\dagger \varphi$.
- Solve $G_{nm} a_m = b_n$ using the Gauss-Jordan method.
- Return the trial vector $\chi_{trial} = \sum_{n=1}^{N} a_n v_n$.

Because χ_i's are nearly linearly dependent there is a numerical instability, and the Gram-Schmidt procedure does not produce a true basis of $span(\chi_i)$ due to round off error. Hoverer $span(v_i)$ is a subspace close enough to the relevant subspace of the old solutions. Because the accumulated error is larger at the late stages of the Gram-Schmitt iteration, the ordering of the χ_i is important. We get better results if we number the χ_i's from the newest to the oldest. This way we quite accurately pick up the new relevant directions, while the old ones, which are less relevant, are computed with more round off error. Although the numerical errors still accumulate in old vectors, those vectors are multiplied by smaller coefficients. In this way we can handle the numeric instability and turn it to our advantage.

This method requires $(N^2 + 5N)/2$ dot products and N $M\chi$ matrix-vector applications, and the storage of $3N$ past pseudo-fermion configurations. There are other implementations of the above concepts which can do the same thing using less memory. But most of them suffer from numerical instabilities. The best we found, and actually used in our production runs, is instead of storing $\chi's$, store $v's$, keeping them in the right order - most recent first. This reduces the memory requirement to the storage of $2N$ pseudo-fermion configurations, and has the same performance as the original method. We checked this method using double precision, and there was no visible improvement by reducing the round off-error.

In Table 3.1, we record the mean number of CG steps required to reach the solution normalized relative to starting with $\chi_{trial} = \varphi$. We used the stopping condition that the normalized squared residual (Eq. 3.1) was smaller than 10^{-14}.

The data (see Table 3.1 and Figures 3-3, 3-4), suggest that the more vectors you keep from the past the better starting residue you achieve, and as a result the fewer CG steps you need in order to converge to a given accuracy. Furthermore the number of CG steps is decreasing as δt is decreasing.

N=0	1.00	1.00	1.00	1.00	1.00	1.00	1.00
N=1	0.73	0.80	0.84	0.85	0.85	0.86	0.90
N=2	0.52	0.67	0.72	0.73	0.76	0.78	0.84
N=3	0.30	0.55	0.63	0.67	0.70	0.72	0.82
N=4	0.21	0.44	0.55	0.59	0.63	0.67	0.79
N=5	0.18	0.33	0.45	0.51	0.56	0.61	0.76
N=6	0.18	0.32	0.38	0.43	0.49	0.53	0.72
N=7	0.17	0.30	0.38	0.42	0.45	0.48	0.68
N=8	0.17	0.30	0.37	0.41	0.42	0.47	0.66
N=9	0.15	0.28	0.36	0.39	0.44	0.47	0.63
N=10	0.15	0.29	0.36	0.39	0.42	0.45	0.61
N=11	0.15	0.27	0.35	0.39	0.42	0.45	0.62
δt	0.002	0.005	0.007	0.008	0.009	0.010	0.015

Table 3.1: Number of CG steps needed to converge to the solution from minimum residual extrapolation. The table is normalized with respect to N = 0 (i.e. no extrapolation, $\chi_{trial} = \varphi$). The statistical errors are of the order of 10%.

As explained above, the performance of the method given in "computational time" CT is summarized in Table 3.2, and in Figure 3-5. It is interesting to note how CT depends on the number of the vectors retained. Unlike the number of CG steps, CT does not decrease as δt becomes smaller. Moreover, there is a range of δt where the performance is relatively insensitive to δt. As you decrease the step size, you gained back almost the same performance by reducing proportionally the number of CG steps required for convergence in each step. As discussed earlier, smaller step sizes have the extra advantage of improving the acceptance rate.

3.4 Comparison with the Polynomial Extrapolation Method

The polynomial extrapolation method has the advantage that it requires very little computational effort, just a local sum on each lattice point with fixed coefficients given once and for all by Equation (3.4), costing less than a single CG step. For a polynomial of order N, the only storage requirement is for the previous N pseudo-fermion configurations, $\chi(t_i)$.

Note also that the coefficients of the Minimal Residual Extrapolation method (Eq. 3.3), which *a priori* are generic complex numbers, are in fact very close to coefficients in the polynomial extrapolation (Eq. 3.4) for the first few orders. One way to understand this coincidence is to observe that

N=0	5.86	2.34	1.66	1.46	1.30	1.17	0.78
N=1	4.29	1.88	1.40	1.24	1.10	1.00	0.70
N=2	3.03	1.57	1.20	1.07	0.98	0.91	0.66
N=3	1.78	1.29	1.06	0.96	0.91	0.84	0.64
N=4	1.23	1.02	0.92	0.86	0.82	0.78	0.62
N=5	1.08	0.78	0.74	0.74	0.73	0.71	0.59
N=6	1.04	0.74	0.64	0.63	0.63	0.62	0.56
N=7	0.99	0.71	0.63	0.61	0.58	0.57	0.53
N=8	0.98	0.69	0.62	0.60	0.57	0.55	0.51
N=9	0.88	0.67	0.60	0.58	0.57	0.54	0.49
N=10	0.90	0.67	0.60	0.57	0.55	0.53	0.48
N=11	0.88	0.64	0.59	0.57	0.54	0.53	0.48
δt	0.002	0.005	0.007	0.008	0.009	0.010	0.015

Table 3.2: Computational Time (CT) starting from minimum residual extrapolation. The table is normalized to 1 at $\delta t = 0.010$ for $N = 1$ (or $\chi_{trial} = \chi_{last}$).

for a smooth evolution the determination of coefficients c_i,

$$\chi_{trial} = c_1\chi(t_1) + c_2\chi(t_2) + \cdots + c_N\chi(t_N), \tag{3.8}$$

by a polynomial fit is **equivalent** to fixing the coefficients by making a Taylor expansion of each term, $\chi(t_n)$, in $t_n = n\delta t$ canceling all contributions to $O(\delta t^N)$. To prove this we solve the constraints

$$\sum_{n=1}^{N}(t_i)^{n-1}y_n = \chi(t_i), \tag{3.9}$$

for a polynomial fit $y(t) = y_1 + ty_2 + \cdots + t^{N-1}y_N$ to find $y_1 = \chi_{trial}$. We then show that $y_1 = \sum_i c_i\chi(t_i)$ when we enforce the Taylor series constraints,

$$\sum_{i=1}^{N}(t_i)^{n-1}c_i = \delta_{1,n}. \tag{3.10}$$

From this exercise, we conclude that the success of the polynomial fit probably results from the local convergence of the power series MD in time. On the other hand, while the polynomial actually deteriorates at high order, the minimal residual continues to improve. At the extrapolation order of $N = 5 - 6$ the deterioration becomes appreciable.

In conclusion, the increased efficiency of the MRE method is worth the extra computational effort (equivalent to $N/2$ CG steps). With a Polynomial Extrapolation, one gets poor results on a few exceptional lattices, whereas the MRE method is a more robust estimator, which implements a kind of self-tuned extrapolation, never doing worse than a polynomial fit of the same order.

3.5 Conclusions

Let us summarize our perspective on the problem of accelerating a time sequence of conjugate gradient inverters. The chronological method has offered significant performance improvement, but at the same time there is a tantalizing need for further improvement. Just by starting from the old solution, the residue is reduced by 4 orders of magnitude to order 10^{-4}, relative to the residue with $\chi = \varphi$. Then if we look at Figure 3-6, we see that the residue is reduce further by 5 orders of magnitude using 10 additional past solution vectors in our Minimal Residual Extrapolation method. Finally, the first 10 CG iterations accounts for another 2 orders of magnitude. However, accurate reversibility ultimately requires us to reduce the residue by another 4 more orders. This last 4 orders of magnitude takes several hundred additional vectors in the CG iterations. We are both intrigued and frustrated by the observation that we can reduce the residue by 11 orders of magnitude in a 20-30 dimensional vector space, but then the standard CG iterations requires hundreds of additional search directions to accomplish the remaining 4 orders of magnitude needed to satisfy adequately the reversibility constraint. It is tempting to hope that further improvements can be made on this last 4 orders of magnitude.

We have emphasized the analytic properties of $\chi(t)$ because we believe it may suggest ways to understand and further improve the chronological method. Our success so far is probably due to the slow evolution of the low eigenvalues of the Dirac operator. Thus we are in essence taking advantage of "critical slowing down" in the HMC algorithm to accelerate the Dirac inverter. However, there may well be other vectors (besides the final solution) in the nearby past iteration that can better exploit the slow evolution of our matrix. For example, the last CG routine, which is closest in time to our present inverter, itself generates many A conjugate search vectors, that may be more useful than the older solutions exploited in our MRE method.

In conclusion, we found that the procedure presented here, and in particular the Minimal Residual Extrapolation method reduces by a factor of about 2-3 the mean number of conjugate gradient iterations required to move in phase space for a fixed molecular dynamics time. This is achieved with negligible extra computational time, at the expense of memory. Consequently when sufficient memory is available for storing the past solutions the Dirac inverter, our Chronological Inversion method certainly provides one more useful trick for more efficiently generating full QCD configurations.

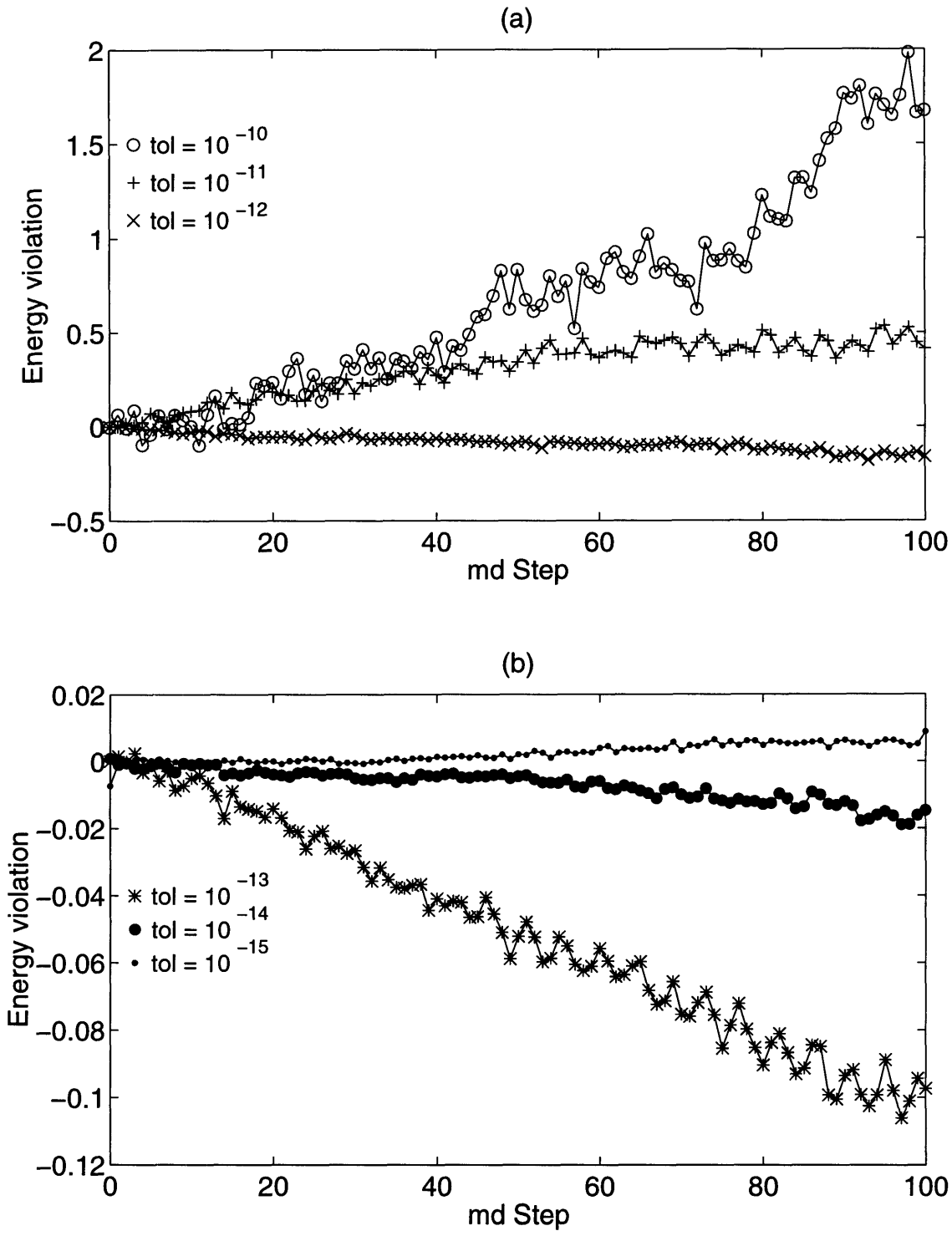

Figure 3-2: Reversibility violations for the energy differences in a forward-backward trajectory with the stopping conditions $R = 10^{-10}, 10^{-11}, 10^{-12}$ (a) and $R = 10^{-13}, 10^{-14}, 10^{-15}$ (b).

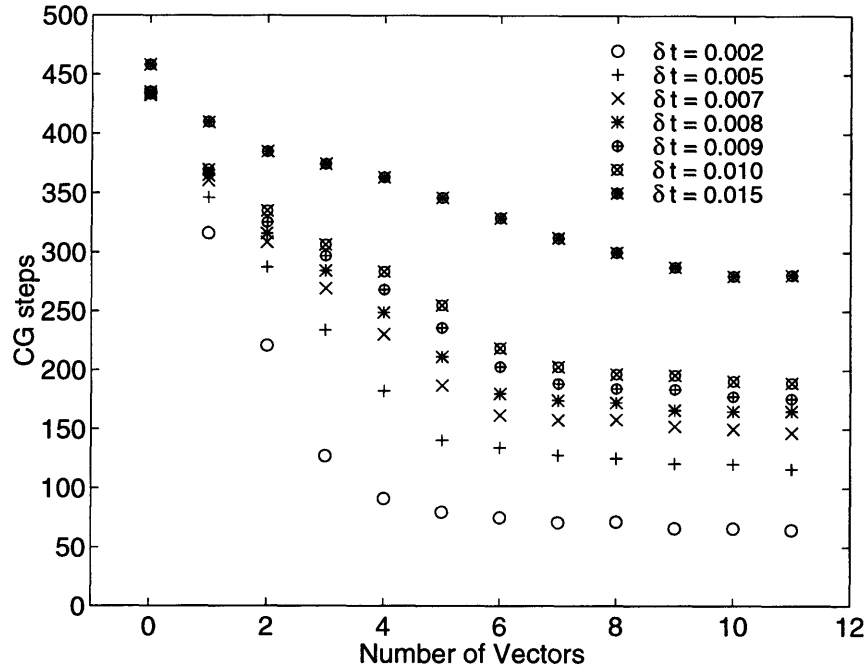

Figure 3-3: Number of CG steps vs. number of vectors for various values of δt for the Minimal Residual Extrapolation method.

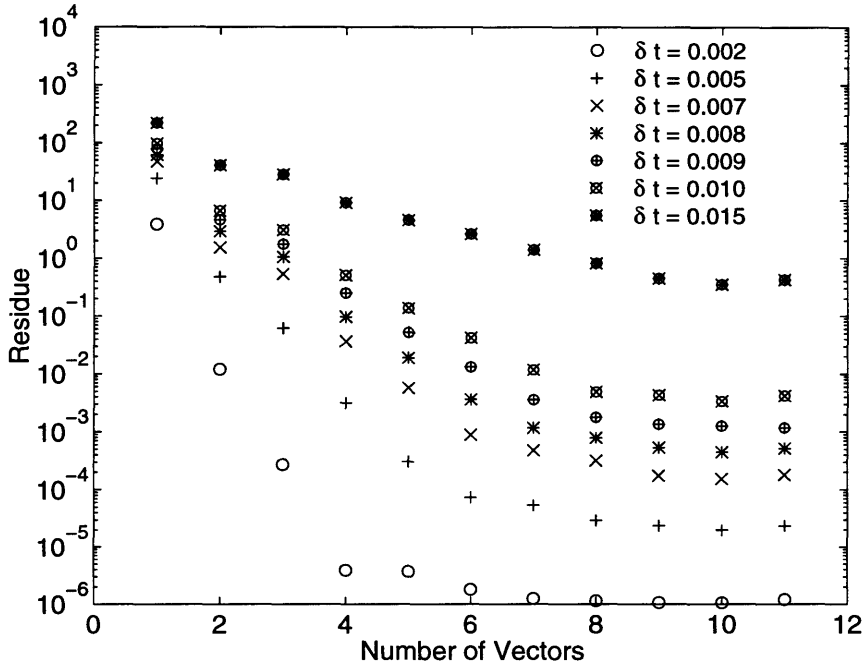

Figure 3-4: Starting residue vs. number of vectors for various values of δt for the Minimal Residual Extrapolation method.

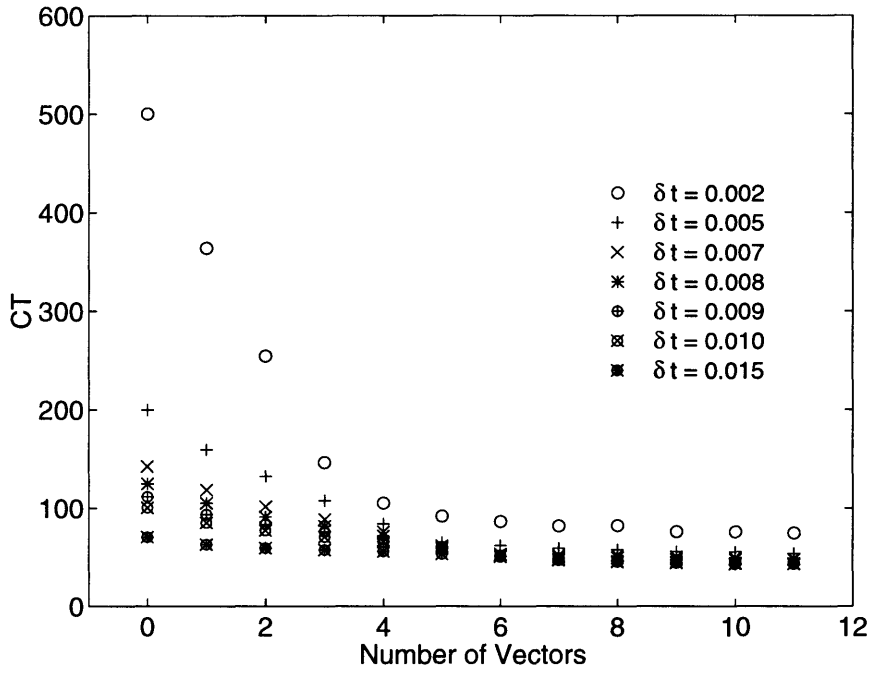

Figure 3-5: CT vs. number of vectors for various values of δt for the Minimal Residual Extrapolation method.

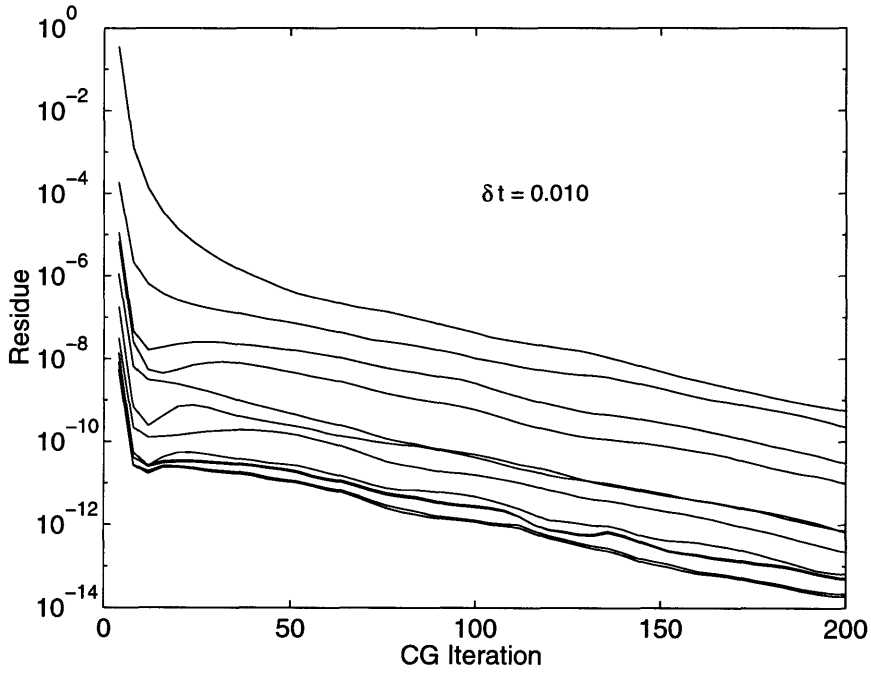

Figure 3-6: Example of convergence of the CG as a function of the number of extrapolation vectors for a single lattice and $\delta t = 0.010$. The method used was Minimal Residual Extrapolation and the number of vectors varies from 0 (top line) to 11 (bottom line).

Chapter 4

Instantons on the Lattice

This Chapter begins by reviewing the properties of instantons in continuum theories and models of instanton dynamics (Sec. 4.1). The variant of cooling algorithm (Sec. 4.2) and the procedures to extract the instanton content of a lattice configuration (Sec. 4.3) are then described and the effects of the relaxation procedure on isolated instantons (Sec. 4.4) are studied. Finally, the instanton distribution of the lattices including dynamical quarks is determined [59].

4.1 Properties of Instantons in Continuum Field Theory

This section reviews the results of instanton models in QCD. The comparison of those results with the lattice calculation provides both validation of the instanton models from first principles and a check for the lattice algorithms.

Instantons are non-trivial localized solutions to the classical equation of motion of the Euclidean Yang–Mills theory. Using the transformation

$$S = \frac{1}{4}\int d^4x F^a_{\mu\nu}F^a_{\mu\nu} = \frac{1}{4}\int d^4x \left[\frac{1}{2}(F^a_{\mu\nu} \pm \tilde{F}^a_{\mu\nu})^2 \mp F^a_{\mu\nu}\tilde{F}^a_{\mu\nu}\right], \tag{4.1}$$

the action is related to the topological charge

$$Q = \frac{1}{32\pi^2}\int d^4x F^a_{\mu\nu}\tilde{F}^a_{\mu\nu}. \tag{4.2}$$

Since the topological charge is integer for the fields going to zero (pure gauge) at infinity, the minimum of the action in a given topological charge sector is reached for (anti-)self-dual fields

$$\tilde{F}^a_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}F^a_{\rho\sigma} = \pm F^a_{\mu\nu} \tag{4.3}$$

A single instanton solution in SU(2) gauge theory, first obtained by A. Polyakov and coworkers

[7, 57] can be written in the following form (due to t'Hooft):

$$A^a_\mu(x) = \Omega_{a,b}\bar{\eta}^b_{\mu\nu}\partial_\nu \log(1 + \frac{\rho^2}{(x-c)^2}), \tag{4.4}$$

where $\Omega_{a,b}$ is a constant rotation matrix in an adjoint representation, c is the instanton position and ρ is the instanton size. The tensors $\bar{\eta}^b_{\mu\nu}$ are defined in Appendix A. This field has a topological charge $Q_I = 1$ and action $S_I = \frac{8\pi^2}{g^2}$, the anti-instanton field with $Q_A = -1$ and $S_A = S_I$ can be obtained by replacing symbols $\bar{\eta}$ with η in the formula above. The instanton solution was later generalized to arbitrary numbers of instantons and arbitrary gauge group [1]. Since the topological charge can not be changed by a small perturbation of fields, the instantons are stable against field perturbation except for a change in parameters. The action of the instantons does not depend on the instanton parameters and small variations of the instanton field (Eq. 4.4) with respect to the parameters are zero modes of the action.

Since the instantons are the classical solutions to the equation of motion, they enter the semi-classical expansion of the partition function along with the true vacuum state. The configurations with well-separated instantons, although not solutions of the classical equations of motion, also contribute to the partition function giving the instanton gas model. In this model all instantons are considered infinitely far apart, the interaction between them is ignored and the partition function can be written as a sum over instantons

$$\mathcal{Z} = \sum_{\text{instantons}} e^{-S_I} \int D' A e^{-S'[A]} \tag{4.5}$$

with all quantum fluctuations factorized in $S'[A]$ and the functional integration is done on the modes orthogonal to instanton zero modes. The sum over all instantons can be written as an integral over the instanton zero mode parameters: positions of the centers $c^{(i)}$, sizes $\rho^{(i)}$ and orientations $\omega^{(i)}$

$$\mathcal{Z} = \sum_{N_i} \left[\prod_i^{N_i} \int \frac{d\omega^{(i)} d^4 c^{(i)} d\rho}{\rho^5} e^{-\frac{8\pi^2}{g^2}} \right] \mathcal{Z}_{\mathcal{N}_i}{}' \tag{4.6}$$

Taking into account quantum corrections, the differential number of instantons is [14, 79]

$$dN_i = \frac{0.466 \exp(-1.679 N_c)}{(N_c-1)!(N_c-2)!} \left(\frac{8\pi^2}{g^2}\right)^{2N_c} \exp\left(-\frac{8\pi^2}{g^2(\rho^2)}\right) \frac{d^4 c d\rho}{\rho^5} \tag{4.7}$$

The instanton distribution in a gas (Eq. 4.7) naively diverges at small ρ and converges at large ρ due to the factor ρ^{-5}. When running of the coupling constant $g(\rho)$ is taken into account, the instanton distribution changes to $\frac{dN_I}{dV d\rho} \sim \rho^{b_0 - 5}$ with b_0 the first coefficient from the beta function (Eq. 2.2). In pure gauge theory $\frac{dN_I}{dV d\rho} \sim \rho^6$, in the theory with two light fermion flavours $\frac{dN_I}{dV d\rho} \sim \rho^{14/3}$. The instanton distribution is finite in the region of small ρ but diverges for large ρ. This divergence

is an artefact of the dilute gas approximation and presumably is removed by non-perturbative effects and the instanton interaction which grows with increasing ρ.

More sophisticated models take into account the instanton interactions [14, 67, 68, 69, 70, 71, 73]. Drawing the analogy with gas condensing into liquid when the relative strength of the interaction increases, these models are called the instanton liquid models. The starting point for such models is an approximate expression for the gauge fields between interacting instantons. The simplest expression is the sum *ansatz*

$$A^a_\mu(x) = \sum_{i\in I,A} \Omega^i_{a,b} A^{(i)\,a}_\mu(x - c_i), \tag{4.8}$$

but it suffers from singularities. The ratio *ansatz*

$$A^a_\mu(x) = \frac{\sum_{i\in I} \Omega^{(i)}_{a,b} \bar{\eta}^b_{\mu\nu} \frac{x^{(i)}_\nu}{\rho_i^2} + \sum_{j\in A} \Omega^{(j)}_{a,b} \eta^b_{\mu\nu} \frac{x^{(j)}_\nu}{\rho_j^2}}{1 + \sum_{i\in I} \frac{x^2_{(i)}}{\rho_i^2} + \sum_{j\in A} \frac{x^2_{(j)}}{\rho_j^2}}, \text{ where } x^{(i)} = x - c^{(i)}, \tag{4.9}$$

does not have singularities. The best approximation is given by the streamline solution which is a minimum of an action in all but one directions in the configuration space [3, 4, 62, 82]. The discussion of instanton effects is continued in Chapter 5, where the instanton effects in the quark propagation are investigated.

4.2 Relaxation Algorithm

The instantons are not directly visible in the lattice configurations obtained by Monte-Carlo algorithms. The quantum noise is usually several orders of magnitude higher than the instanton field strength. To separate the instanton content from the quantum noise, the lattice configuration must be made smoother by some procedure. Equivalently, the lattice observables must be redefined to a smooth combination of point-like original observables. The standard smoothing procedure is cooling, when each link on the lattice is changed so that the local action reaches the minimum. After several steps of cooling, the quantum noise is dramatically reduced while the solutions to the classical equation of motion, instantons, do not change much. The fermionic methods provide an example of the second way of averaging the quantum noise (see Chapter 5).

In order to extract the instanton content of the lattice configurations efficiently on a parallel computer, a variant of the cooling method has been developed by discretizing the relaxation equation in the extra "time" coordinate τ. Symbolically, the relaxation equation is

$$\frac{\mathrm{d}U}{\mathrm{d}\tau} = -\frac{\delta S}{\delta U}. \tag{4.10}$$

Since the matrices U must stay unitary during the relaxation process, the Equation (4.10) must be

modified to take into account the unitarity constraint. The SU(3) manifold in the vicinity of the matrix U can be parameterized with a Hermitian traceless matrix F by $U[F] = \exp(iF)U$. The parameters in F, for example, coefficients in a λ-matrix expansion, have no constraints and the relaxation equation (4.10) is an ordinary relaxation equation in a linear space of parameters F

$$\frac{\mathrm{d}F}{\mathrm{d}\tau} = -\frac{\delta S}{\delta F} \tag{4.11}$$

with an initial condition $F_0 = 0$. When this equation is discretized for a small step of size ϵ in the "relaxation time" τ, the evolution step can be written as

$$U' \leftarrow exp\left(-i\epsilon\frac{\delta S}{\delta F}\right) U \tag{4.12}$$

and *all* links on the lattice are updated simultaneously. The action gradient is determined by varying the action

$$S[U'] = \mathrm{Re\ Tr}\, U'U_s = S[U] + \mathrm{Re\ Tr}\ [iFUU_s] \tag{4.13}$$

where U_s is a generalized "staple", a sum of the products of all loops entering the action. The final form of the functional derivative is fixed by the requirement that the derivative must be a Hermitian traceless matrix,

$$2\frac{\delta S}{\delta F} = UU_s - (UU_s)^\dagger - \frac{2}{N_c}\mathrm{Im\ Tr}\,[UU_s]. \tag{4.14}$$

For large values of ϵ this algorithm is unstable but in the limit $\epsilon \to 0$ the action evolution curve converges to the solution of the relaxation equation (4.10) (Fig. 4-2). The curves for the topological charge do not converge, which indicates some instabilities inherent to our system. For SU(3) gauge fields, the value $\epsilon = 0.025$ gives fast and stable relaxation and this value has been used in all measurements. Comparing the evolution of the action for our relaxation method with $\epsilon = 0.025$ and the "standard" Cabibbo-Marinari cooling, one observes that one cooling step is approximately equivalent to 4-8 relaxation steps and that the cooling histories for the action are very similar (see Figure 4-1). The effective number of relaxation steps corresponding to one cooling steps depends on the configuration so it is impossible to establish exact correspondence.

The relaxation algorithm has a well-defined scaling properties in a continuum limit. In this limit Equation (4.10) transforms to

$$\frac{\mathrm{d}A}{\mathrm{d}\tau} = -\frac{\delta S}{\delta A} \tag{4.15}$$

with the continuum gauge field A. When the field A is rescaled by the factor α, the evolution time τ must be rescaled by α^2 to keep the form of the equation. In comparison, the ordinary cooling algorithm does not have such a property.

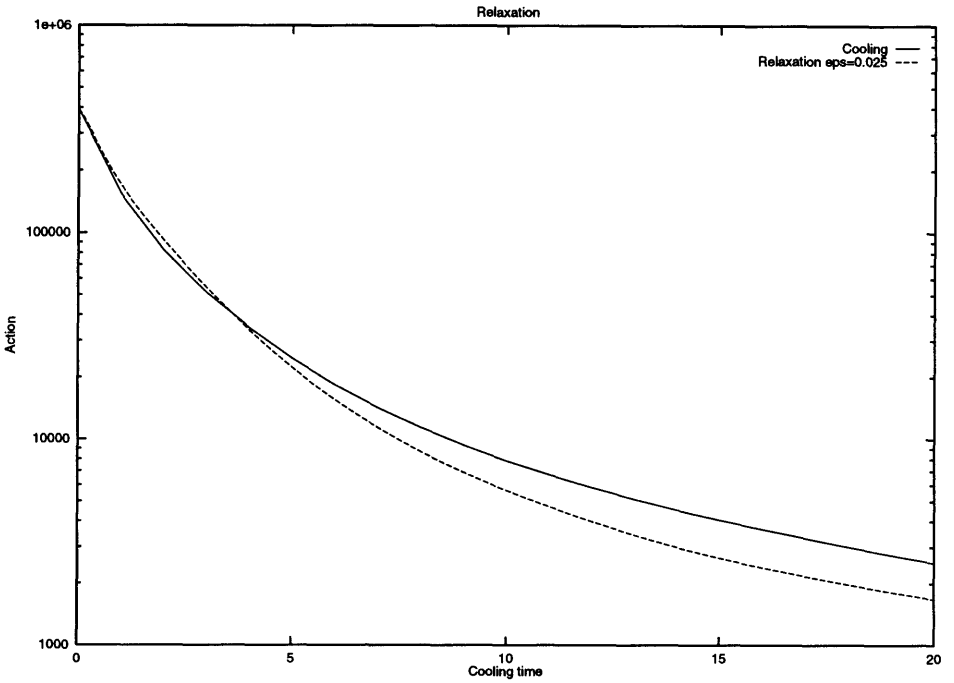

Figure 4-1: Comparison of the action evolution under cooling and relaxation algorithms for a random lattice. On the horizontal axis are the number of the cooling steps and relaxation time multiplied by 5.

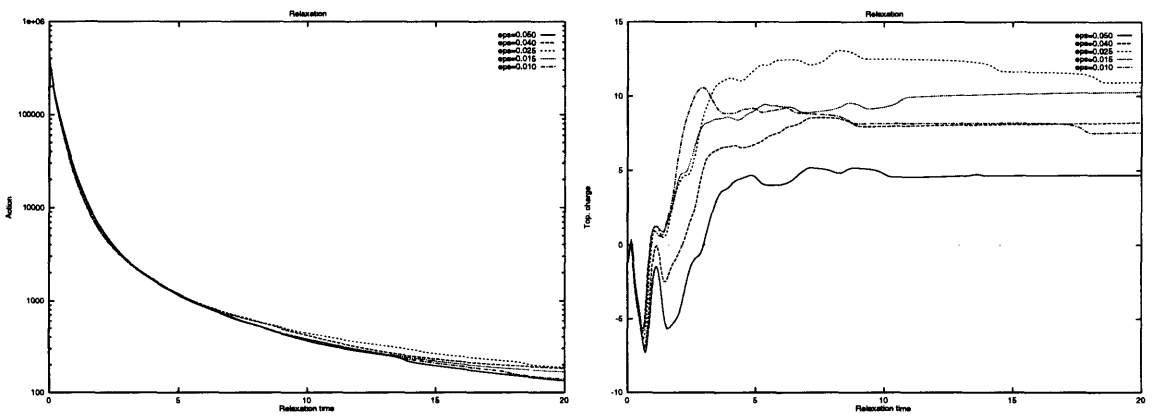

Figure 4-2: Comparison of the relaxation algorithm for different values of ϵ.

4.3 Identifying Instantons

We used the following procedure to identify instantons in our configurations. First, we consider all peaks in the action and topological charge density as candidates for the instantons. With an initial value of the instanton position c_0 (peak position) and instanton size $\rho_0 = (48/S(c_0))^{1/4}$ (with the normalization $S(x) \to a^4 F\tilde{F}$ in the continuum limit), we perform a least squares fit to the action and topological charge densities of a classical continuum instanton,

$$S_0(x|c,\rho) = \frac{48\rho^4}{((x-c)^2+\rho^2)^4}. \tag{4.16}$$

as a function of c, ρ. (The details of the fitting are described below.) If the fitting process converges and the final values for the instanton position c and size ρ are within a pre-defined range close to the initial values, we record the identified instanton. If one of the conditions above is not satisfied, the candidate peak in the action or topological charge is discarded.

In the simplest approximation we considered the instantons as well-separated objects and used the fit function (Eq. 4.16) for each instanton in turn. In the next approximation we used the combined fit function

$$S_1(\{c_i,\rho_i\}) = \sum_i S_0(x|c_i,\rho_i) \tag{4.17}$$

In this function the effects of the interaction are completely lost. We have also tried a more complicated fit function like the ratio *ansatz* (Eq. 4.9), but the improvements to the fit were only marginal and the amount of additional information extracted from the distribution did not justify the extra work. In all cases studied in this work, the algorithm fails when two instantons become close, approximately on the distance $d \approx \rho$. An obvious generalization would be to fit instanton–anti-instanton pairs with a streamline solution when d is close to ρ.

The instanton action density falls off rapidly with the distance from the center and is screened by other instantons and remaining noise on distances $d \sim \rho$. There is little information to gain by fitting outside this region so the weights defining the least-squares fit are cut off beyond this distance. The following expression for the χ^2 is used:

$$\chi^2 = \sum_x w(x-x_0)\left(S(x) - S_0(x_0,\rho_0)\right)^2, \tag{4.18}$$

where S_0 is the action density in the selected approximation and the weight is

$$w(r) = \frac{1+e^{-ab}}{1+e^{a(\frac{r}{\rho_0}-b)}}. \tag{4.19}$$

This weight goes smoothly from 1 to zero in the region of the width ρ_0/a centered at $b\rho_0$ where ρ_0 is the original estimate of the instanton size. The results in this work use the values $a = 5$ and $b = 1$.

The fit is rejected if the values of the parameters go outside the fit region $x^i \in [x_0^i - b\rho_0, x_0^i + b\rho_0], \rho \in [0.7\rho_0, 1.5\rho_0]$ and $\rho \in [\rho_{min}, \rho_{max}]$. From studying the evolution of a single instanton (Sec. 4.4), the values $\rho_{min} = 2.5a$ and $\rho_{max} = 0.75L_x$ were determined. The algorithm works slightly better when the action density in Equation (4.18) is replaced by the topological charge density.

Although there is some degree of arbitrariness in selecting the parameters in the algorithm, we were able to identify around 50% of the peaks in the action distribution as instantons with the remaining peaks corresponding to small instantons, closely bound instanton–anti-instanton pairs or other excitations.

4.4 Isolated Instantons

Several effects can distort the instanton distribution during the relaxation process. The smallest instantons disappear by falling through the mesh. The larger instantons may shrink because of interaction with periodic images or may disappear by instanton–anti-instanton pair annihilation. We separated these effects by studying the evolution of semi-classical instanton configurations. We discretized the t'Hooft solution (Eq. 4.4) by integrating $U = P\int \exp A dx$ along each link and used sum *ansatz* (Eq. 4.8) to produce instanton configuration. When the instanton positions are selected close to the center of hypercube, the gauge field on the boundary is small and we ignored the dislocation caused by forcing the field to be periodic. After a few relaxation steps the configuration becomes very close to a semi-classical solution of the lattice equation of motion.

First, we investigated the shrinking of a single isolated instanton for a variety of instanton sizes and two lattice sizes, 16^4 and 24^4.

In Figure 4-3 we show the evolution of a single isolated instanton during relaxation using the lowest-order Wilson action. During relaxation the trajectories begin at the value of ρ specified in the key and the size shrinks monotonically until the instantons shrinks to $\rho = 2.5a$ at which point they "fall through the lattice". The reason for this behaviour is that the first order term in the Wilson action expansion in a^2/ρ^2 is negative and the relaxation process forces instantons to shrink. Since relaxation only adiabatically changes the instanton parameters, we can reconstruct the $S_I(\rho)$ dependence by taking the envelope of evolution histories for different runs. Since there is no plateau in the action, instanton-like semiclassical solutions do not exist. Hence, the Wilson action is completely inadequate for instanton study.

Because of the deficiency of the Wilson action, we studied also a first order improved action [35],

$$S_{\text{imp}} = \frac{4}{3}\text{Tr}(1 - W_{1\times1}) - \frac{1}{48}\text{Tr}(1 - W_{2\times2}). \tag{4.20}$$

The coefficients in the improved action were chosen so that the zeroth-order term in a^2 is the same as in Wilson action, the first order term is zero and the second order term is positive. As Figure

4-4 shows, the small instantons are much more stable with the improved action [22, 23, 24, 25, 35]. Adding more terms to the action could further stabilize instantons [22, 24, 25, 35] but the first-order improved action was simple and already sufficient for our studies.

For large instantons the evolution does not depend appreciably on the choice of the action but the boundary effects are substantial. Unlike other groups [35] we decided to study the instantons in a periodic box. Ultimately we wanted to study the dynamical configurations produced with periodic boundary conditions and changing boundary conditions for relaxation appears inconsistent. In a periodic box, the mirror images put pressure on the instanton and always force it to shrink. It may be possible to counteract this pressure, described by a ρ^2/L^2 term in the action, by adding terms increasing the action for small instantons, but this kind of fine-tuning appears unnatural to us. Since the dependence of action on ρ^2/L^2 is so strong, extra terms could stabilize the action of the instantons only in small region in ρ, distorting the rest of the distribution. Choosing bigger lattices significantly reduces finite-volume corrections as shown in Figure 4-5.

With the improved action, Figures 4-4 and 4-5 show that the minimal size of an instanton that can be sustained on the lattice is approximately $2.8a$. Furthermore, the disappearance of the $\rho = 3.0a$ instanton on the 16^4 lattice in Fig. 4-4 is a boundary effect, since it survives much longer on the 24^4 lattice in Fig. 4-5. Figure 4-5 shows that the improved action is superior in several important respects. The action for the smallest instantons is much closer to the continuum result, the slope is significantly lower and the second derivative is positive. As discussed, the remaining increase of S with ρ is associated with finite–volume effects.

Since we are ultimately interested in on the order of 100 relaxation steps at most, we conclude that with the improved action and significantly larger lattices, the artificial loss of single instantons in cooling is negligible.

Next, we studied the evolution of an instanton–anti-instanton (I-A) pair. Figure 4-7 shows the history in a typical case: instantons of initial sizes $\rho_I = \rho_A = 6.0a$ with the separation between centers $s = 12.0a$. On the small lattice 16^4, the boundary effects are much stronger than the interaction between objects and they shrink in place before they have a chance to interact. On the larger lattices, the instanton and anti-instanton attract each other and annihilate. The finite volume effects have stabilized on the largest lattices. When the instanton and anti-instantons have different initial sizes, they first seem to equalize sizes and then disappear (Fig. 4-8)

We have also checked that the discretization effects are negligible by comparing the evolution of two appropriately rescaled configurations. Figure 4-6 shows the evolution of the two physically equivalent configurations on 16^4 and 24^4 lattices. The evolution curves almost coincide, which means that the finite lattice spacing corrections are small in this case.

The "interaction" time during which the instanton and anti-instanton move toward each other is much less then the "shrinking" time and thus the individual instantons in a pair do not shrink

significantly. In the case of real dynamical configurations, we expect the interactions of instantons to dominate their shrinking and hence we believe the conventional use of periodic boundary conditions will not introduce serious finite volume distortions.

4.5 Instanton Distribution in Quenched and Full QCD

We have studied 21 full QCD configurations for $\beta = 5.5$, 2 flavors of Wilson fermions and $\kappa = 0.160$. As measured by LANL group [39, 41], this choice of parameters corresponds to the lattice spacing determined from the pion decay constant $a(f_\pi) = 0.11$fm and pion mass $m_\pi \approx 630$MeV. Our own measurements (see Sec. 5.7), although limited in statistics, agree with those estimates. Configurations were separated by 50 HMC trajectories with average length 50 steps of size $\epsilon = 0.01$. Those configurations have the same value of $\kappa_c = 0.16145$ as quenched QCD for $\beta = 5.858$ (See Ref. [39, 41]). To compare quenched and full QCD in roughly the same physical region, we have generated 23 quenched configurations for $\beta = 5.85$ separated by 500 heat-bath iterations. Although one can not strictly define a string tension with dynamical quarks, we have also checked that the slope of the quark–anti-quark potential (measured by the Creutz ratios) in both cases is comparable in the range of 3-5 lattice units.

Figure 4-9 shows the distribution of instantons in sizes after 20, 30 and 50 relaxation steps. Since our method for identifying instantons is not reliable before 20 steps, we can not determine the early evolutions of the distribution, but after 20 the distribution certainly changes dramatically with relaxation. Given that the number of instantons is $O(200)$ or larger in a 16^4 box, the average distance between instantons is about 4 lattice units, comparable to the average instanton size. Under these conditions the interaction between instantons is very strong and consequently the effects of I-A pair annihilation are much stronger than boundary effects.

Since the instanton interaction is so strong that the number of instantons drops very fast with relaxation, it is not possible to choose the "optimal" number of relaxation steps for measurements of instanton distributions. This result is in sharp contrast to the results of the other groups [25, 23, 24] which also measured instanton distributions on a different set of lattices. There are several possible explanation for discrepancies.

First, the degree of cooling used in [24] is much higher than in the present studies. The $20 - 50$ relaxation steps in this work, correspond to < 10 cooling steps, whereas the other group started with 20 cooling steps and continued to 300. There was not much additional erosion of distribution after 100 relaxation steps in this work either but the distribution of instantons has already changed dramatically after so many steps and its relation to the original distribution is not clear.

Second, [24] used the SU(2) gauge group. The dynamics of instantons in SU(3) and SU(2) gauge groups may be different due to a different number of degrees of freedom.

Lastly, it is clear that the configurations obtained by Monte-Carlo procedures have close, strongly interacting instanton–anti-instanton pairs. At some region of instanton sizes ρ the dynamics of QCD can no longer be fully described by instantons and the transition from instanton dynamics to perturbative dynamics is not well-defined. Although small instanton pairs add to the total instanton distribution, their influence to physical values may be small [69, 67, 68, 70, 71, 73]. For the same reason, the total density of instantons, $\approx 1\text{fm}^{-4}$ in the instanton liquid model [73] is much less then the density we observed on a lattice $\sim 200 * (0.11\text{fm} * 16)^{-4} = 21\text{fm}^{-4}$.

Since the erosion of the instanton distribution with cooling should be the same for the quenched and unquenched samples, we believe the fact that the distributions shown in Fig. 4-9 are essentially identical within errors provides strong evidence that the physical instanton distributions are very similar in quenched and full QCD, at least at this sea quark mass, which is of the order of 100MeV.

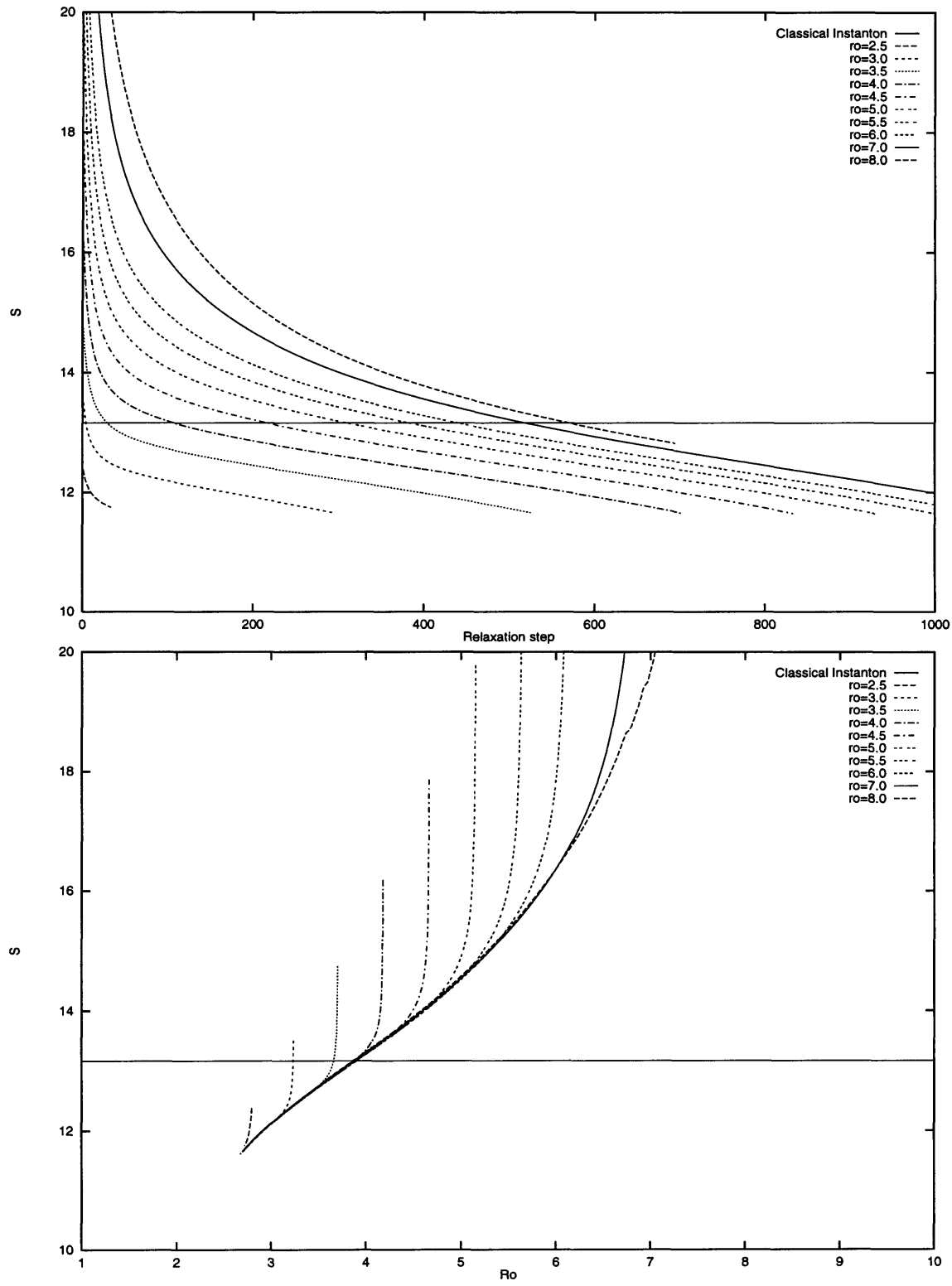

Figure 4-3: Evolution of single isolated instantons of different sizes on a 16^4 lattice with Wilson action

On the upper graph, the action S is shown as a function of the number of relaxation steps for each initial value of ρ shown in the key. In the lower graph, S is plotted versus the instanton size ρ for the same cases.

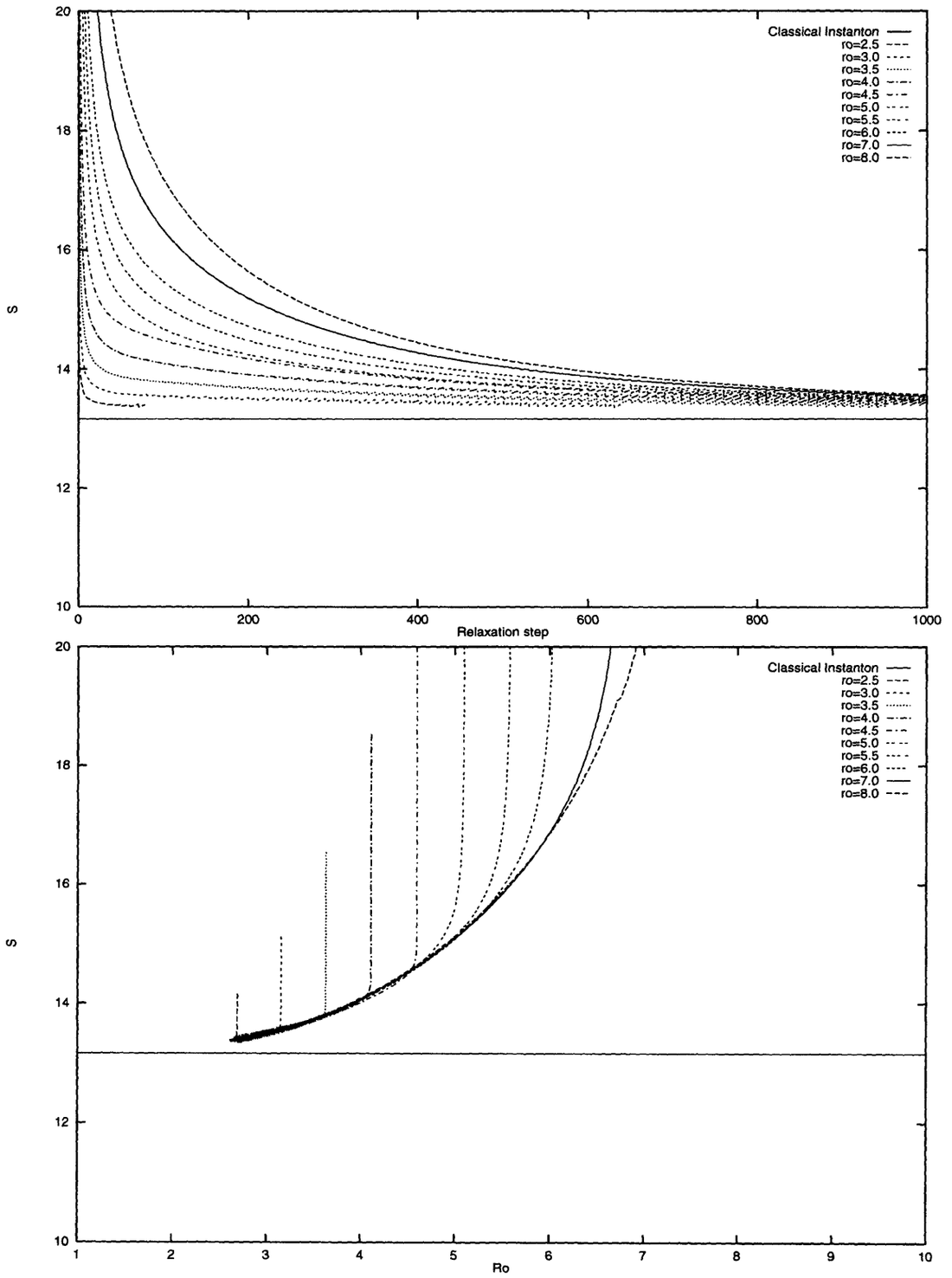

Figure 4-4: Evolution of single isolated instantons of different sizes on a 16^4 lattice with improved action

On the upper graph, the action S is shown as a function of the number of relaxation steps for each initial value of ρ shown in the key. In the lower graph, S is plotted versus the instanton size ρ for the same cases.

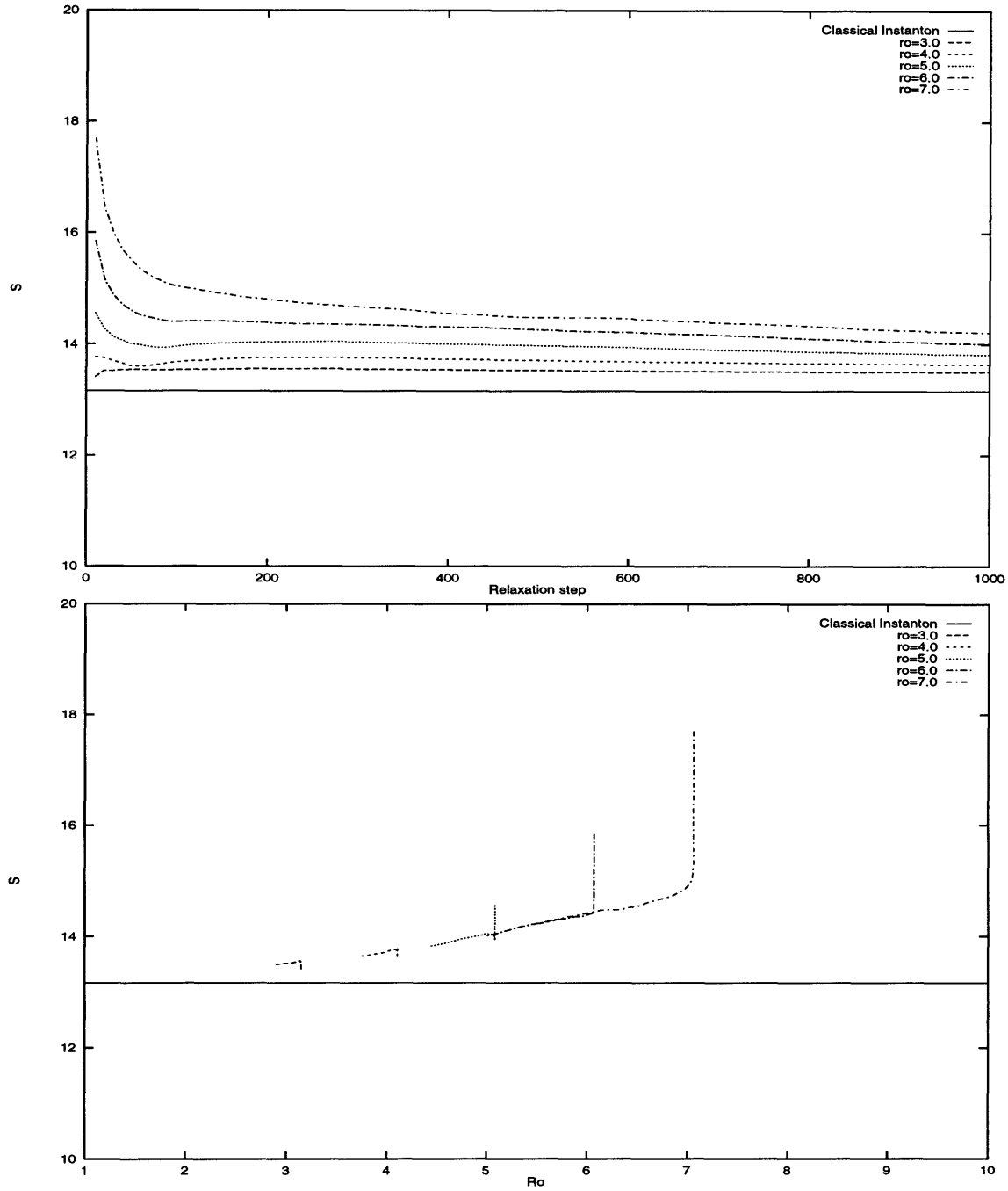

Figure 4-5: Evolution of single isolated instantons of different sizes on a 24^4 lattice with improved action.
On the upper graph, the action S is shown as a function of the number of relaxation steps for each initial value of ρ shown in the key. In the lower graph, S is plotted versus the instanton size ρ for the same cases. Note that the termination of curves is due to stopping of the relaxation, not disappearance of instantons. The curves for the action on the upper graph do not go uniformly down due to accumulation of the round-off errors in the single-precision arithmetic.

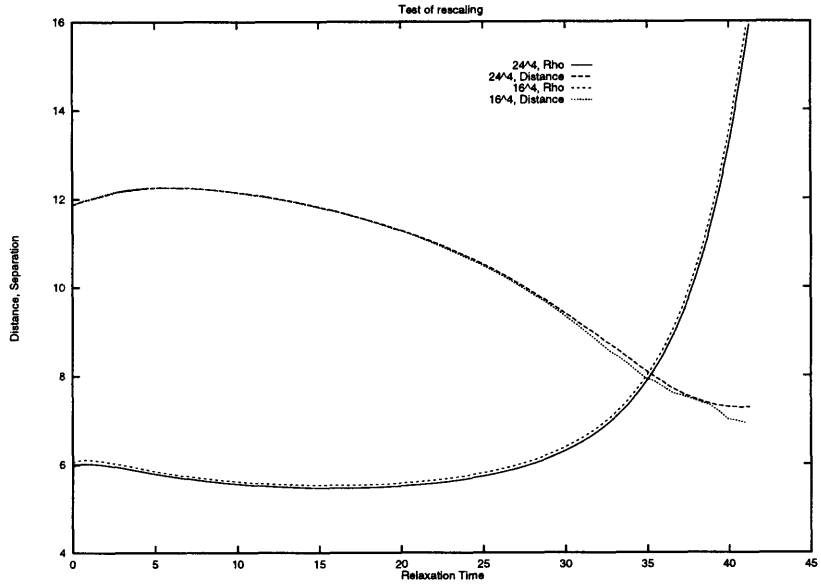

Figure 4-6: Test of rescaling. The configuration on the 24^4 lattice is an A-I pair of sizes $\rho_A = \rho_I = 6$ separated by $s = 12$. The configuration on the 16^4 lattice is an A-I pair of sizes $\rho_A = \rho_I = 4$ separated by $s = 8$, with the same relative orientation as in the first case, and the relaxation time is multiplied by 2.25.

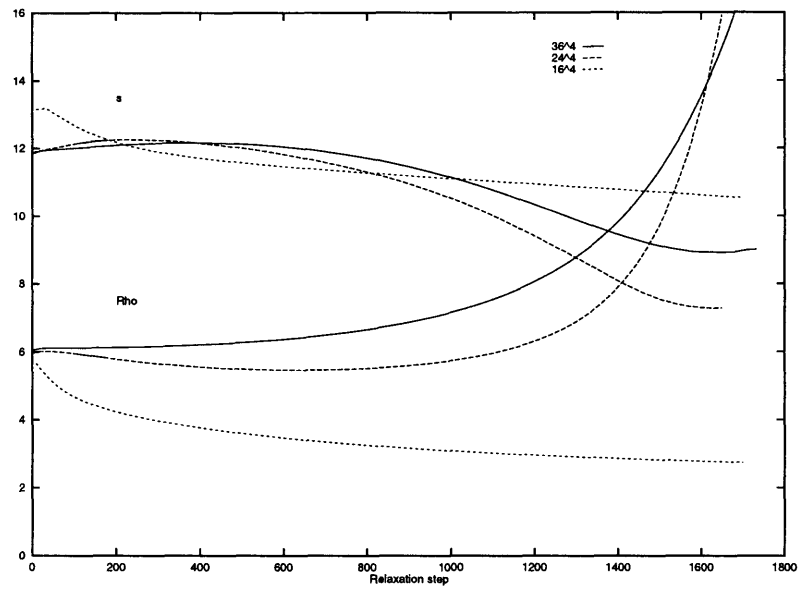

Figure 4-7: Evolution of an I-A pair under relaxation on lattices of different size. The upper curves give the I-A separation s and the lower the sizes ρ. The "data" for the 36^4 lattice are obtained by a re-scaling from a 24^4 lattice.

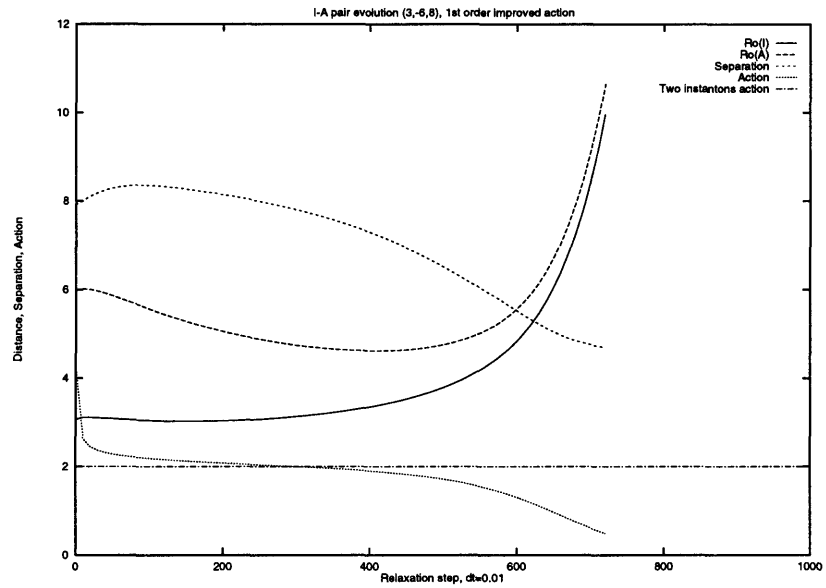

Figure 4-8: Evolution of an asymmetric I-A pair under relaxation on a 16^4 lattice. The initial configuration had $\rho_A = 3$, $\rho_I = 6$ and separation $s = 8$.

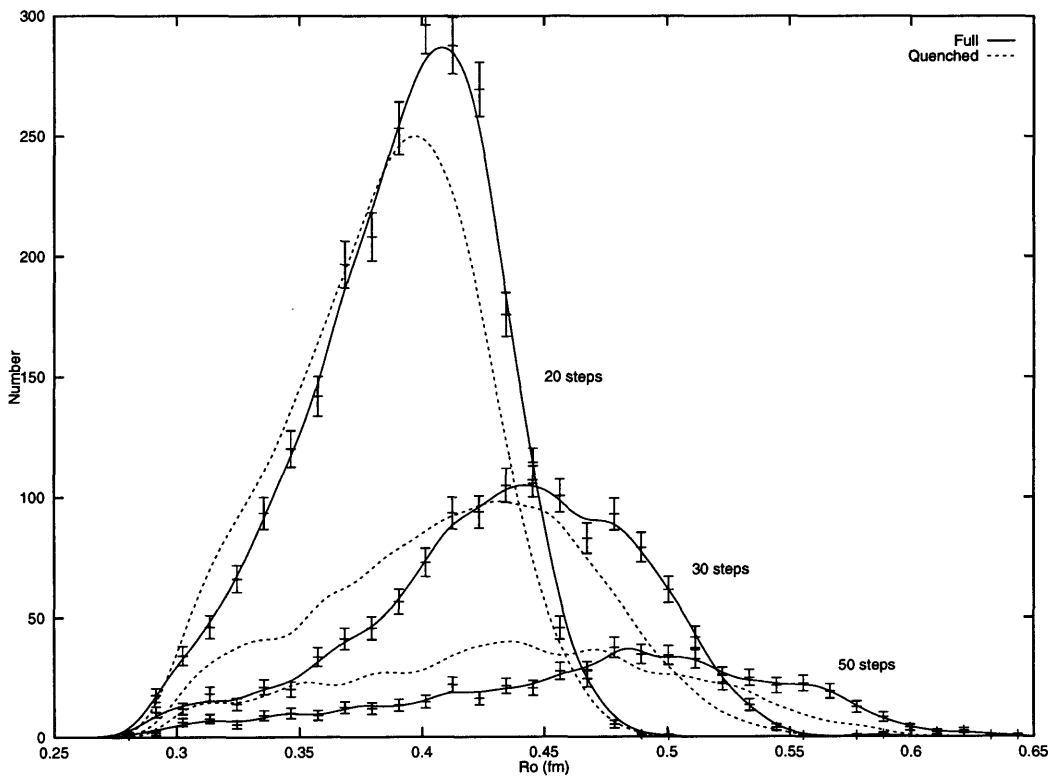

Figure 4-9: The instanton distributions in quenched and full QCD. The smooth curve is the measured distribution convoluted with a Gaussian curve of width 0.01fm. Error bars are estimated by the jackknife method, and are comparable for both cases.

Chapter 5

Study of the Spectrum of the Dirac Operator on the Lattice

5.1 Properties of Wilson-Dirac Operator

This section summarizes the general properties of the Wilson-Dirac operator D defined in Eq. (2.6) or the even-odd preconditioned version M defined in Eq. (B.3).

- The operator D is neither Hermitian nor anti-hermitian but satisfies the following commutation relations with γ_5:

$$D = \gamma_5 D^\dagger \gamma_5 \tag{5.1}$$

- All eigenvalues of D come in complex-conjugate pairs or are real. As a result the partition function for full QCD (Eq. 2.10) is also real.

- If ψ_i is an eigenvector of D then $\gamma_5\psi_i$ is an eigenvector of $D^\dagger$ with the same eigenvalue. The eigenvectors of D and $D^\dagger$ are related by a simple γ_5 multiplication.

- In general, two distinct eigenvectors are not orthogonal but if ψ_i and ψ_j are eigenvectors of M corresponding to eigenvalues λ_i and λ_j

$$\langle\psi_i|\gamma_5|\psi_j\rangle \neq 0 \text{ only if } \lambda_i = \lambda_j^\dagger \tag{5.2}$$

  Eigenvectors with real eigenvalues are γ_5–orthogonal to the rest of a spectrum and complex eigenvectors are γ_5–null vectors. We ignore the unnatural possibilities of exact zero eigenvalues and use relation above as equivalence (It is assumes that for each ψ_i there is another vector ψ_j such that their γ_5–product is non-zero).

- Using γ_5 orthogonality, the unity decomposition can be written as

$$1 = \sum_i \frac{|\psi_i\rangle \langle\psi_{i'}| \gamma_5}{\langle\psi_{i'}|\gamma_5|\psi_i\rangle} \tag{5.3}$$

$$= \sum_i \frac{\gamma_5 |\psi_i\rangle \langle\psi_{i'}|}{\langle\psi_{i'}|\gamma_5|\psi_i\rangle} \tag{5.4}$$

5.2 Free Fermion Spectrum

Figure 5-1 shows the calculated spectrum of the Wilson-Dirac operator for free fermions. In this case the imaginary axis is the norm of momentum and the real axis is the effective mass. There are 5 low-momentum regions where 16 doublers live. The region with the lowest effective mass is the physical region in the continuum limit. It is clear from the analytic expression (Eq. 2.8) that the first set of doublers (around the real point 0.5) has momentum $(\pi, 0, 0, 0)$ and behaves like $(-1)^x$. There are four such doublers, one for each of the directions on the lattice. Due to the Wilson term, those modes have acquired the mass $\approx 0.5a$. When the lattice spacing $a \to 0$, this mass goes to infinity and doublers disappear from the spectrum. Only the eigenvectors in the physical region are smooth in a continuum limit and the effective mass in this region (Eq. 2.7) is finite in continuum limit. The true continuum spectrum is, of course, $m + i|p|$ — the straight line parallel to imaginary axis and deviations from this line come from the Wilson term. For subsequent reference it is useful to denote the real part of the lowest eigenvalue $\lambda_{\min} = 1 - 8\kappa$ (see Eq. 2.8).

One more complication arises from the finite lattice size. When periodic boundary conditions are imposed on all directions, the spectrum has a zero-momentum constant mode. This mode is not normalizable in the infinite-volume limit but complicates the study of the low-energy modes. There is no such mode when anti-periodic boundary conditions are imposed in time direction and this boundary conditions were used in this work. These boundary conditions are also necessary for the correct formulation of the transfer matrix and definition of the continuum limit.

The free fermion spectrum is highly degenerate. The lowest mode has $p_{(0,1/2)} = (0, 0, 0, \pm\frac{\pi}{L_t})$ and there are $8N_c$ such modes. The next mode is $p_{1,1/2} = (\pm\frac{2\pi}{L_x}, 0, 0, \pm\frac{\pi}{L_t})$ with multiplicity $48N_c$. Since the algorithm used in this work is limited to the 128 lowest eigenmodes on 16^4 lattice, those two modes are the only ones found numerically but the eigenvalues for those modes are barely visible on the edge of Figure 5-1.

5.3 Spectrum in the Field of One Instanton

The first non-trivial example is a spectrum in the field of one instanton. This spectrum is not known analytically but it has exactly one normalizable zero mode [79]. In a more general case, the index theorem [2] specifies that the difference between the numbers of left and right zero modes is equal

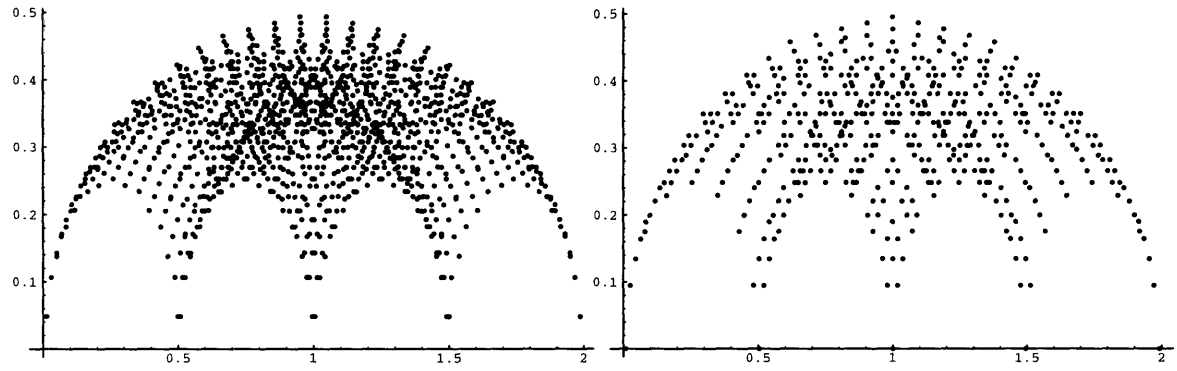

Figure 5-1: The spectrum of free Wilson Fermions with anti-symmetric and symmetric boundary conditions. Only the upper half of the spectrum shown, the lower half is symmetric. The hopping parameter is $\kappa = 0.124$, which gives the fermion mass $m_q = 0.032$ and $\lambda_{min} = 0.008$.

to the topological charge. The zero mode for an instanton at $x = 0$ with the size ρ has the form

$$\Psi_0(x)_{s,\alpha} = u_{s,\alpha} \frac{\sqrt{2}}{\pi} \frac{\rho}{(x^2 + \rho^2)^{3/2}} \tag{5.5}$$

with some constant spinor $u_{s,\alpha}$ such that $u^\dagger u = 1$. This (continuum) mode has definite parity, left for instantons and right for anti-instantons. On the lattice this mode can have a non-zero right component due to the Wilson term in the operator and finite–volume corrections [75]. Figure 5-2 shows the low-λ part of the spectrum in the field of a single instanton with $\rho = 4.5$ on 16^4 lattice. The spectrum shows a clear "zero" mode (labeled "0"), two more localized modes and four clusters of other modes. These modes are discussed in detail below.

The zero mode is not exactly at zero on the lattice but is shifted due to the Wilson term. The shift can be calculated in first order of perturbation theory (in the Wilson coefficient r in Eq. 2.6) with the result

$$\mathrm{Re}\lambda_0 = \langle \Psi_0 | D | \Psi_0 \rangle = \frac{3\kappa a}{\rho^2}, \tag{5.6}$$

which gives a reasonable estimate for the position of "zero-mode" eigenvalue. The Figure 5-4 shows the calculated positions of zero mode eigenvalues and the prediction of the shift Eq. (5.6). This shift has important implications on the dynamics of fermions in the instanton background. First, zero modes in the continuum have exactly the same eigenvalue and mix under arbitrarily small interactions. On the lattice, the "zero" eigenvalues for different size instantons are separated and

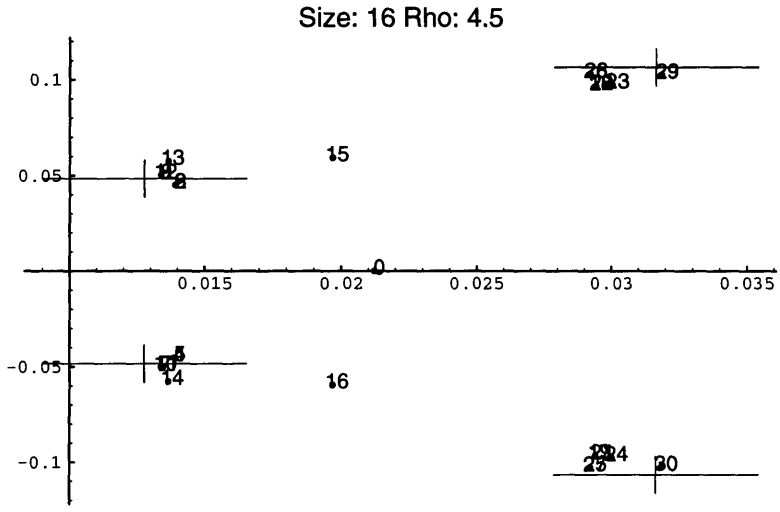

Figure 5-2: The lowest 32 modes of the Wilson-Dirac operator in the field on instanton $\rho = 4.5$ on a 16^4 lattice. The hopping parameter $\kappa = 0.124$ corresponds to mass $m = 0.032$, $\lambda_{min} = 0.008$. Crosses denote the positions of eigenvalues for free fermions with this κ.

sufficiently small interactions do not mix corresponding eigenmodes (see Section 5.4). Due to the Wilson term, the fermion mass is renormalized to $m_{eff} = m + 3ra/2\rho^2$.

The zero mode is not the only low-energy mode on the lattice. There are also plane-wave like modes which have magnitude $\sim 1/\sqrt{V}$. Those modes belong to the continuum part of the spectrum in continuum but in finite box the zero mode and continuum modes are not well-separated. The instanton zero mode mixes with plane-wave modes and produces several more low-lying localized modes. The lowest mode for the free fermions has momentum $p_{(0,1/2)} = (0, 0, 0, \pm\frac{\pi}{L_t})$ and there are $8N_c$ such modes which differ by color and spinor orientation. Since the instanton zero mode has a definite and constant color and spinor orientation, it mixes strongly with only two of the $p_{(0,1/2)}$ modes; the mixing with other $p_{(0,1/2)}$ modes is suppressed. In Figure 5-2 the strongly mixed modes are labeled "15" and "16" and weakly-mixed modes are clustered around free value $0.0127653 + 0.0483824i$ (see Eq. 2.8). Since the plane-wave modes are smoother then the instanton zero mode, the effects of the Wilson term in the mixed modes are smaller. In Figure 5-2 the real part of the mixed mode eigenvalues "15" and "16" is clearly between the values for instanton zero mode and the cluster of weakly-mixed modes. As Figure 5-3 shows, the positions of eigenvalues of

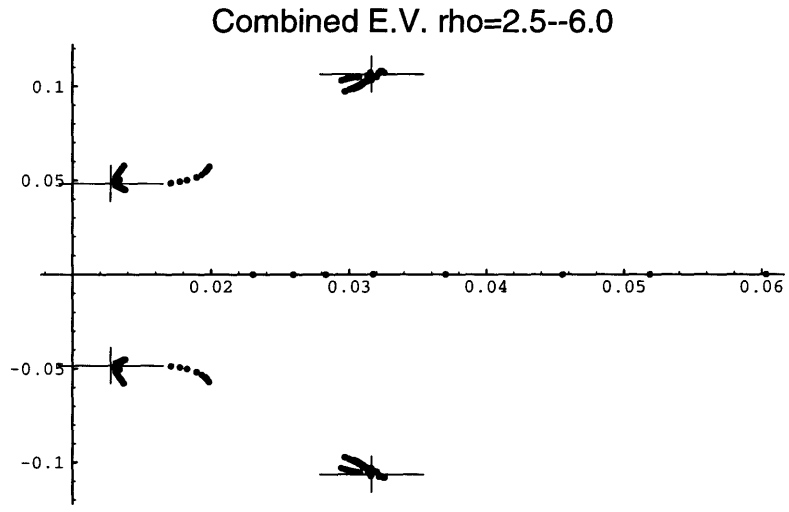

Figure 5-3: The lowest 32 modes of the Wilson-Dirac operator in the field of instantons of different sizes. Crosses denote the positions of eigenvalues for free fermions with this κ.

the weakly-mixed modes do not depend much on the instanton size, the position of strongly-mixed modes have some dependence on instanton size and the position of the zero modes have a strong dependence described by Eq. (5.6). The other cluster of eigenvalues, visible in the Figures 5-2 and 5-3, corresponds to the modes with momentum $p_{(1,1/2)} = (\pm\frac{2\pi}{L_x}, 0, 0, \pm\frac{\pi}{L_t})$ and (free) eigenvalues $\lambda(p_{(1,1/2)}) = 0.0316431 + 0.106527i$.

Since the instanton mode has a definite chirality, it can be separated from non-chiral plane waves. Figure 5-5 shows selected modes for an isolated instanton with $\rho = 4.5$ on the 16^4 lattice. On the right is the square of the right components, in the center is the square of the full eigenvector and to the left is the square of the left components. The top row is the instanton zero mode. It is dominated by left components and has a very small right component. The maximum value for the norm of this mode is $\Psi_0^2(0) = \frac{2}{\pi^2\rho^4} = 0.0005$ (See Eq. 5.5), close to the lattice value. The next row is the strongly-mixed mode labeled "15". The left component has the same shape as the zero mode but is approximately 2 times smaller in magnitude. The plane wave contribution is not visible on the left graph but dominates on the right graph for right component. Mode "16" is similar and is not displayed. The third row shows the mode "13" which has momentum $p_{(0,1/2)} = (0, 0, 0, \pm\frac{\pi}{L})$ and is weakly mixed with the instanton. The left and right components have clear sine profiles with

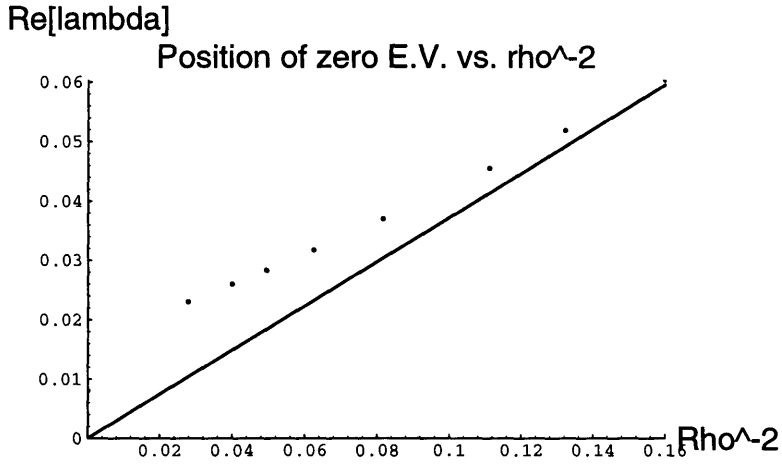

Figure 5-4: Position of the "zero-mode" eigenvalue in the field of one instanton vs. the instanton size. The dots correspond to calculated zero modes in the field of lattice instantons and the straight line is the perturbative prediction Eq. 5.6.

almost constant sum characteristic of the plane wave. The amplitude of this mode agrees with the prediction for the plane wave on the lattice $\Psi^2 = 1/V = 1.53 * 10^{-5}$. The last row is one of the modes with momentum $p_{(1,1/2)} = (0, 0, \frac{2\pi}{L_z}, \frac{\pi}{L_t})$. It has a sine profile in both Z and T directions with some deviations around the position of the instanton.

Figure 5-6 shows the similar modes for the same size ($\rho = 4.5$) instanton on the 24^4 lattice. Due to memory limitations, it was not practical to calculate as many modes in this case. The relative strength of the plane wave modes is smaller in this case ($\Psi^2 = 3 * 10^{-6}$) while the strength of the instanton zero mode is approximately the same. By comparing the middle row with the top row one can see that the mixing coefficient squared is about $1/4$, half the value for 16^4 lattice. Since in perturbation theory we expect the mixing amplitude to vary as $1/L$, the square should decrease by $4/9$ in going from 16^4 to 24^4 lattice. Thus, the scaling is as expected and in the infinite volume limit the mixing coefficient should go to zero.

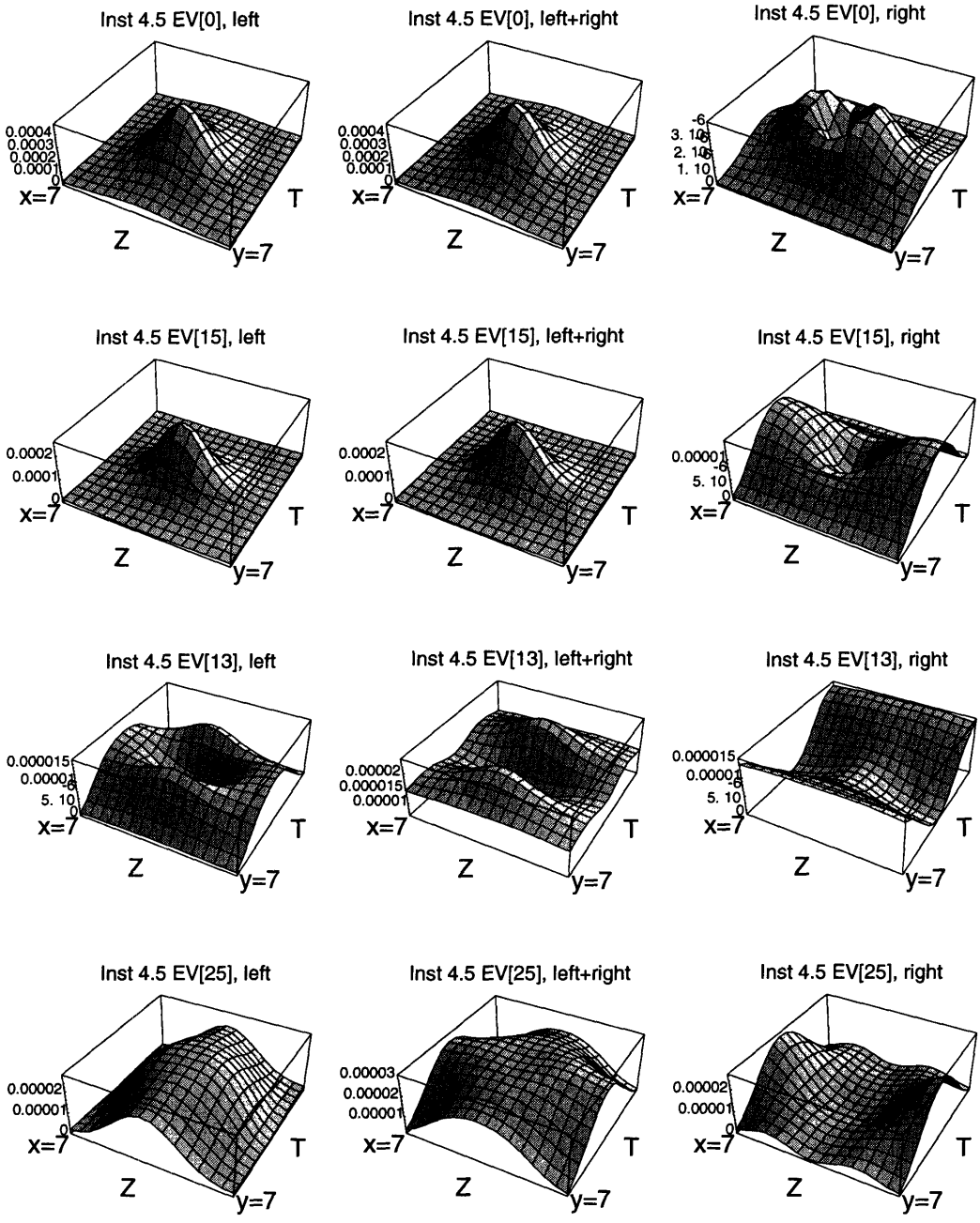

Figure 5-5: Chiral projections of selected eigenmodes of an instanton with $\rho = 4.5$ on a 16^4 lattice. Eigenvalue (EV) numbers correspond to the labels on the eigenvalue plot Fig. 5-2. The top row is the instanton zero mode labeled "0", the next row is the mixture of the instanton zero mode and the first Matsubara mode $p_{(0,1/2)}$ labeled "15", then the Matsubara mode weakly mixed with the instanton zero mode labeled "13" and on the bottom row is the mode $p_{(1,1/2)}$.

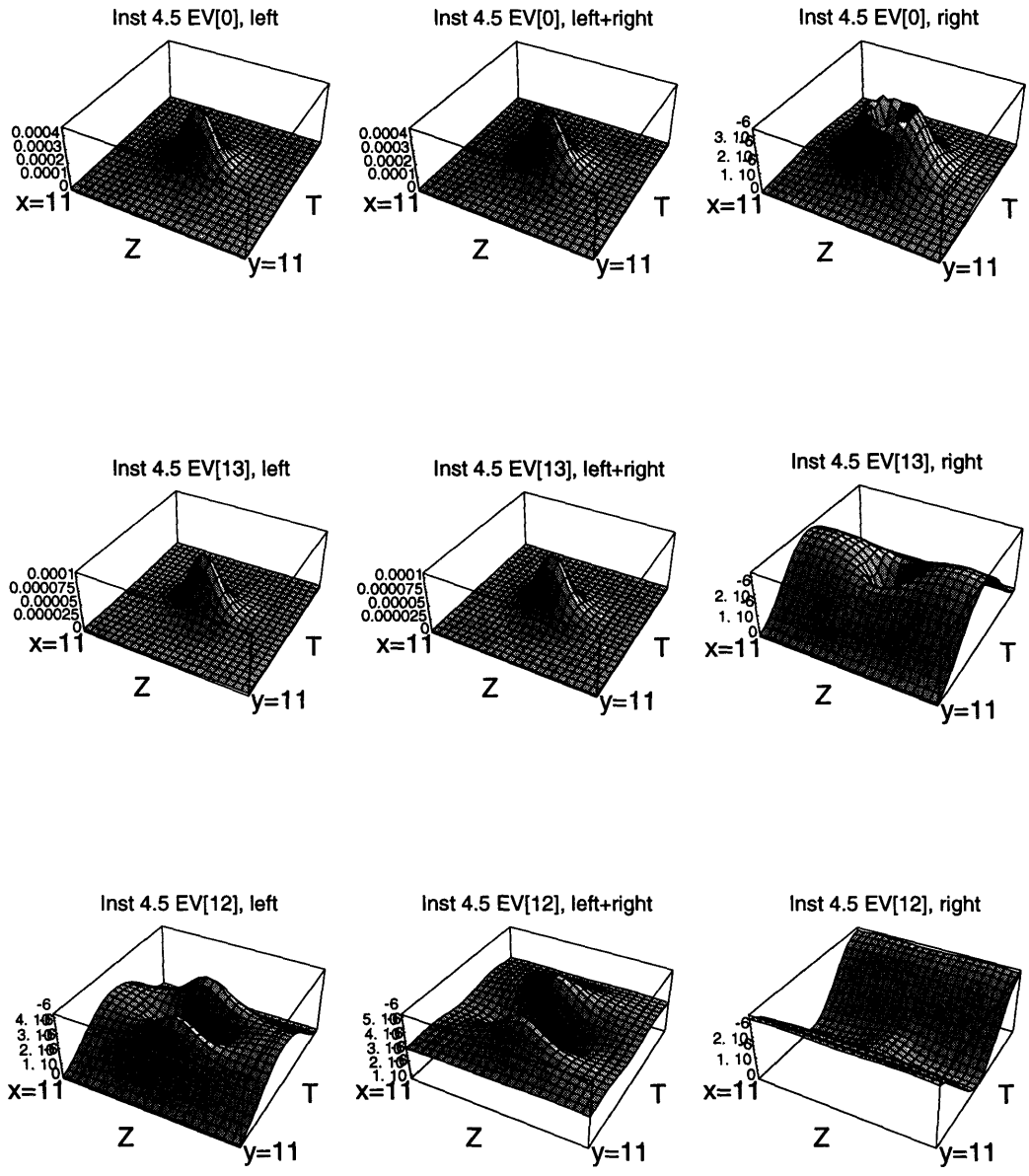

Figure 5-6: Chiral projections of selected eigenmodes of an instanton with $\rho = 4.5$ on a 24^4 lattice. On the top row is the instanton zero mode labeled "0", the next row is the mixture of the instanton zero mode and the first Matsubara mode $p_{(0,1/2)}$ labeled "13", then the Matsubara mode weakly mixed with the instanton zero mode labeled "12". Note that the numbering of eigenvalues differs from those on Figure 5-5 for the 16^4 lattice.

5.4 Effects of Instanton Interaction in the Spectrum

Having understood the modes for a single instanton, we considered the spectrum of the Dirac operator in the field of an instanton–anti-instanton pair. There are no exact zero modes in this case even in the continuum limit but, when the pair is well separated and the interaction is small, zero modes localized at each instanton do not change much and the true eigenmode must be a mix of two zero modes.

Consider the subspace spanned by zero modes which is called the Zero-Mode Zone [64, 72, 73, 83]. The Dirac operator in this subspace has the following structure in the first approximation:

$$D \begin{pmatrix} \psi_I \\ \psi_A \end{pmatrix} = \begin{pmatrix} \lambda_I & T_{IA} \\ T_{AI} & \lambda_A \end{pmatrix} \begin{pmatrix} \psi_I \\ \psi_A \end{pmatrix}, \tag{5.7}$$

where λ_I, λ_A are the eigenvalues for a single zero mode and $T_{IA} = \langle \psi_A | D | \psi_I \rangle$ is the overlap matrix element. The overlap elements T_{II} and T_{AA} vanish because of chirality. The overlap element T_{IA} can be calculated in either the sum (Eq. 4.8) or streamline [62] *ansatz* for the gauge field appearing in D.

In the continuum, $\lambda_I = \lambda_A = 0$ and $T_{AI} = -T_{IA}^\dagger$ so that the ZMZ Dirac matrix is simplified to the matrix

$$D_{\text{ZMZ}} = \begin{pmatrix} 0 & T_{IA} \\ -T_{IA}^\dagger & 0 \end{pmatrix} \tag{5.8}$$

with eigenvalues $\lambda_{1,2} = \pm i |T_{IA}|$. As we see, the interaction between instantons produces a pair of complex conjugate eigenvalues.

On the lattice with Wilson fermions, the situation is complicated by the absence of chiral symmetry and the Wilson term in the operator. We can count on $\lambda_{I,A}$ being real but, for unequal instanton sizes, they are no longer equal. The positions of the eigenvalues of the matrix (Eq. 5.7)

$$\lambda_{1,2} = \frac{\lambda_I + \lambda_A}{2} \pm \sqrt{\frac{(\lambda_I - \lambda_A)^2}{4} - |T_{IA}|^2} \tag{5.9}$$

depend on the relative strength of the Wilson term (Eq. 5.6) and overlap matrix elements. Therefore, we can not expect the formation of the clean ZMZ on the lattice. Rather, the ZMZ will consist of the real eigenvalues *and* pairs of complex eigenvalues split by the overlap matrix.

The Figure 5-7 shows the evolution of the spectrum of an I-A pair during relaxation. In addition to smoothing the configuration, relaxation also smoothly evolves the instanton parameters so that the dynamics of the spectrum becomes obvious. When, during cooling, instantons become wider and move closer together, the "zero" eigenvalues move towards each other, collide and split into a complex pair. At this point the instantons can not be considered isolated objects, at least from the

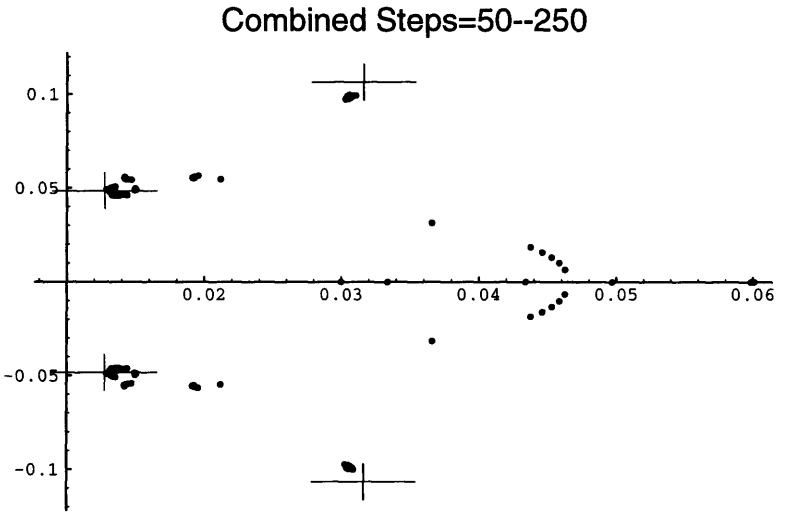

Figure 5-7: Evolution of the spectrum of an I-A pair during relaxation.
The initial sizes of the instanton and anti-instanton were $\rho_I = 2.5$, $\rho_A = 4.0$ and the distance between centers was $|x_A - x_I| = 4.0$. The initial orientations of the pair was chosen arbitrarily. All 32 lowest eigenvalues for the relaxation steps $50 - 250$ are plotted simultaneously. The value of hopping parameter $\kappa = 0.124$ corresponds to mass $m = 0.032$ and $\lambda_{min} = 0.008$. Crosses denote the positions of eigenvalues for free fermions with this κ.

fermion point of view. Eventually all eigenvalues move toward the free spectrum (Eq. 2.8).

5.5 Eigenvectors in Dynamical Fields

The structure of the spectrum of the Wilson-Dirac operator in the dynamical gauge configuration, although not known in general, can be studied using instanton models in the low-energy region [64, 65, 73, 74]. The spectrum has two distinct parts: discrete low-energy modes corresponding to zero modes of the instantons present in the configuration and the continuum of other modes. In the instanton liquid model the quark propagator in the subspace of instanton zero modes is considered separately from the continuum and the effects of the continuum modes are absorbed into renormalization of coupling constants. In this approximation the Wilson-Dirac operator is completely described by a (finite) overlap matrix of zero modes similar to Equation (5.8). Further discussion may be found in ref. [73].

On the lattice the spectrum is complicated by both finite–volume effects and finite–spacing effects. In the continuum, the mixing between instantons is complete — even a small interaction mixes the instanton zero modes. The spectrum is determined by the whole instanton ensemble so

that no instanton can be considered isolated. On the lattice, the Wilson term prevents mixing unless the interaction is strong enough and thus favours the isolated instantons.

Figure 5-8 shows the lowest 128 eigenmodes for a typical configuration and for the same configuration after 100 relaxation steps (see Sec. 4.2). Since the instantons do not change much under cooling, the low-imaginary-part of the spectrum should be similar in those cases. Unfortunately, there is no clear way to separate the instanton modes from low-momenta finite-volume waves. After 100 relaxation steps all but a few eigenvalues come close to the continuum spectrum.

5.6 Using The Lowest Eigenvectors to Accelerate the Conjugate Gradient Algorithm

Having determined the lowest part of the spectrum, it is useful to revisit the problem of accelerating the Conjugate Gradient algorithm (see Chapter 3). This algorithm is used to solve the inversion problem of the Dirac operator

$$D\chi = \psi \tag{5.10}$$

The convergence of the algorithm is determined by the condition number $|\lambda_{min}/\lambda_{max}|$. Since the lowest few eigenmodes were determined exactly (up to machine precision), Equation (5.10) can be projected onto the subspace of zero modes and the complement to it.

$$\begin{cases} P_0 D\chi &= P_0\psi \\ (1-P_0)D\chi &= (1-P_0)\psi \end{cases} \tag{5.11}$$

with a projector P_0 defined as a partial sum in Eqs. (5.3) or (5.4). The first equation can be solved algebraically or using CG in a subspace at negligible cost while the second requires CG for the modified operator $D_0 = (1-P_0)D$. The lowest mode of the D_0 operator has eigenvalue bigger than λ_{min} so we can expect faster convergence. Also, the dependence of the eigenvalue on the mass is strongest for the lowest modes and, after taking them out, the convergence of the algorithm would not significantly depend on mass.

The improvements to CG were studied with the same configuration for which the spectrum was calculated and shown in Figure 5-8. Figure 5-9 shows the time history for the CG with different numbers of eigenvectors subtracted, that is, included in the projector P_0 in Eq. 5.11. The figure shows monotonic improvement in CG convergence with the few first eigenvectors giving the most improvement. The spikes at the end of curves arise because once the residue drops below 10^{-14}, it is necessary to restart the CG to suppress numerical errors arising from the loss of orthogonality in CG.

The Figure 5-10 shows the time histories of CG for different κ and for several numbers of subtracted eigenvectors. For this configuration the critical value of κ, defined by the first real eigenvalue,

is $\kappa_\lambda = 0.16322$ while the critical κ defined by a statistical average over an ensemble of configurations produced with the same parameters was estimated in [9] to be $\kappa_c = 0.1624$ and the critical κ defined by the pion mass on this single configuration $\kappa_c(\pi) = 0.1628(6)$. Note that the numbers above were obtained by extrapolation in the valence quark mass with the sea quark mass fixed by using $\kappa_s = 0.1600$ in HMC. As expected, the convergence of CG is getting worse when κ approaches κ_c while the convergence of CG with subtracted eigenvectors does not depend much on κ.

The improvements in CG convergence come at the cost of applying the projector at each CG step. The real measure of improvement is the overall computer time needed to reach convergence. Figure 5-11 shows the CG convergence as a function of computer time for the same values of parameters as in Fig. 5-10. Subtracting the first 17 eigenvectors gives almost a factor of two improvement in computation cost. Beyond the optimal point, taking more eigenvectors into account does not give more improvement as the cost of additional projections becomes bigger then the gain in the number of iterations.

Restricting ourselves to the lowest 32 eigenvalues and taking the conservative value of 100 iterations needed to find those vectors, the total computation cost to find eigenvalues is approximately equal to 30000 CG iterations. This number can be reduced by making fewer re-orthogonalizations (see Sec. B.2) and by stopping iterations after finding the few lowest modes of interest. The cost of 12000 CG iterations can recovered by reducing the number of iterations required to find one set of 12 propagators (for 3 colors and 4 Dirac indices) by 1000 iterations each. When three or more sets of propagators are to be computed with different (low) masses, the improvements begin to justify the extra cost. The same method should also improve other algorithms like Biconjugate Gradient and Minimal Residue, although those algorithms were not studied in this work.

5.7 Zero-Mode Zone Approximation of Correlation Functions

The effects of instantons on quark propagation can be studied using hadron correlation functions

$$C_h(x,y) = \langle J_h(x) J_h(y) \rangle \tag{5.12}$$

In this work the meson correlation functions for the currents $J_5 = \bar{\psi}\gamma_5\psi$ and $J_\mu = \bar{\psi}\gamma_\mu\psi$ were studied. At small distances ($r < 0.5$fm) the interaction effects are small and correlation functions approach the free massless correlation function $\sim x^{-6}$. At large distance ($r > 1.5$fm) the correlation function is dominated by the lightest particles in the given channel, π and ρ mesons, and falls off like a free particle propagator $\sim x^{-\frac{3}{2}} e^{-m|x-y|}$. The behaviour of the correlation function at intermediate distance $\approx$ 1fm depends on the details of interactions in the given channel [72]. This

region (0.5fm $\leq r \leq$ 1fm) is the region where results of different models can be compared.

The same correlation functions can be computed in the instanton liquid models using the propagator in the subspace of instanton zero modes [60, 63]. Since the propagator in the basis of eigenmodes falls off like $1/\lambda$, the low-λ modes are enhanced in the propagator. If the instanton zero modes are separated by a gap from the rest of a spectrum, the low-λ eigenmodes are completely described by the mixing of instanton zero modes [81, 80, 43, 42] and all other modes are irrelevant at large separations. Even the simple instanton models give a good approximation to the zero-mode zone structure and correlation functions.

The computed low-λ modes were used to approximate the quark propagator by truncating the spectral representation in this subspace

$$D^{-1}(x,y) = \sum_{n=1}^{N} \frac{|\psi_i(x)\rangle \langle\psi_{i'}(y)| \gamma_5}{\langle\psi_{i'}|\gamma_5|\psi_i\rangle \lambda_i} \tag{5.13}$$

(see Eq. 5.3 for precise definition of the conjugate vector index i') and the correlation functions were calculated in this approximation. This section describes the technical details of the calculations, followed by a discussion of results and comparison with other models.

5.7.1 Procedures

The correlation functions on the lattice have both lattice anisotropy corrections and image corrections, which were treated as in reference [16].

The lattice anisotropy effects are the biggest in the region of small distances where the correlation function is well approximated by the free correlation function. In this region the anisotropy effects must also be similar to those in a free propagator. Hence, we define an anisotropy function as the ratio of the free correlation function on the lattice to the continuum expression for a free propagator

$$\langle J_5(0)J_5(x)\rangle = \frac{4}{\pi^4 x^6}, \tag{5.14}$$

$$\langle J_\mu(0)J_\mu(x)\rangle = \frac{8}{\pi^4 x^6}. \tag{5.15}$$

The image corrections in the region $d < 16$ are small on the 64^4 lattice which was used to define the anisotropy function. The resulting correction function is plotted in Figure 5-12. At large distance, the correction function approaches the constant value $f_{corr}(\infty) \approx 2.892 \pm 0.01$ which can be interpreted as a renormalization constant 0.96 for the lattice δ-function times $N_c = 3$.

To take into account the image corrections the interference terms between the image sources are assumed to be small and the periodic correlation function is approximated by the sum of the terms corresponding to infinite volume correlators from each of the images. The cross terms include the sum over Wilson lines looping across the lattice and the average of such terms in dynamical

configurations is small.

The following functional form was used to parameterize the correlation function,

$$f(x) = c(x^{-6} + w^{-6})\frac{1}{1 + x^{3/2}e^{m(x-b)}}, \tag{5.16}$$

and then we fit the sum of $f(x)$ for all first images to the data with free parameters m, w, c, b. This function has the correct form for the correlation function at both small and large distances and joins them smoothly. As a consistency check it was verified that the crossover parameters w and b are within the fitting range [1:16] and that the mass m is close to the correct value. Since the correlation function changes by several orders of magnitude in the region of interest, it was useful to fit the logarithm of it.

The correlation functions in the low-λ subspace have a different general form. The following fitting function was used to approximate the truncated correlation function:

$$f(x) = c(1 + \frac{x^2}{w^2})e^{-mx} \tag{5.17}$$

At large distances this function still falls off exponentially and some small-distance polynomial corrections are introduced to improve the quality of the fit. Since the eigenvectors were computed on even sites only and the propagators on the odd sites were reconstructed after computing the truncated correlation function, the correlation function is not smooth on even and odd sites. This approach introduces a small error in the correlation function but allows significant savings in memory and computer time. It is possible to correct for such lattice artefacts by multiplying the correlation function on the odd sites by a number close to 1, another fit parameter. For actual comparison with the full correlation function, the truncated correlation function must be divided by $f_{corr}(\infty)$.

The following self-consistent fit procedure was used to check the fit bias and to identify cases in which the parametric fit was not reliable. The lattice correlation function, in the same approximation as before, is given by the sum of contribution of images

$$f(x) = \sum_{i \in \text{images}} \tilde{f}(x_i) \tag{5.18}$$

This equation implicitly defines the image–corrected correlation function $\tilde{f}(x)$. This equation was solved iteratively and after each iteration the solution $\tilde{f}(x)$ was smoothed by replacing $\tilde{f}(x)$ with the (log-)linear approximation to it in the region $x \pm w$. This procedure reduces the statistical errors but introduces the bias of the order $\approx \frac{1}{6}\log(f)''(x)w^2$. Since the correlation function is expected to fall off exponentially, the second derivative is small. With $w = 0.25 - 0.5$ the bias is around 1%, much less then the data spread and the correlator agrees with the parametric fit above within errors.

Figure 5-13 shows the correlation function in the ρ-meson channel as calculated on the lattice

and after applying the lattice anisotropy and image corrections. Figure 5-14 shows the fit to the "truncated" correlation function, and Figure 5-15 compares the parametric fit with the self-consistent solution.

5.7.2 Calculation of Correlation Functions.

In this work correlation functions were calculated for three configurations taken from the set used in Chapter 4: one configuration from the full QCD simulation, the same configuration cooled 20 steps, and a quenched configuration with approximately the same lattice spacing.

First, the meson masses must be determined to set the physical scale. The meson masses for the chosen coupling constants were determined by R. Gupta it et. all [39]. According to this reference, $m_\pi a = 0.366(10)$ and $m_\rho a = 0.519(13)$. To determine the masses in this work, a Gaussian source with the width $3.0a$ was created on time slice 0 on which the gauge field was fixed to Coulomb gauge, and the zero-momentum projected correlation functions were fitted to the sum of the propagator and the first image corrections in the region [5:27]. To increase the fitting region the lattice was periodically doubled in the time direction from 16^4 to a $16^3 \times 32$ lattice.

Knowing the pion mass, the critical hopping parameter κ_c and the quark masses in the given configuration are determined by a chiral fit. The spectrum in the low-λ region also defines the parameter κ_λ, the value of κ where the lowest eigenvalue goes to zero. This value must be very close to κ_c but, in general, the relation between the one single lowest eigenvalue and the meson masses is not clear. The results of the fit are summarized in Table 5.1 and Figures 5-16 and 5-17. With the caveat that we have only used a single configuration and the ρ mass is not linear in κ^{-1}, we note that the lattice ρ-meson mass is roughly $m_\rho a \approx 0.44$, corresponding to a lattice spacing $a_\rho^{-1} \approx 1.8\text{GeV}$ or $a \approx 0.11 fm$. The data in this work are not sufficient for a reliable chiral extrapolation of the ρ-meson mass, and we note that the high-statistics extrapolation in [9] gives the value $a_\rho^{-1} \approx 2.1\text{GeV}$ or $a \approx 0.094\text{fm}$.

Configuration	κ	$m_\pi a$	$m_\rho a$	$m_q a$	κ_c	κ_λ
Full, uncooled	0.1600	0.3507(5)	0.458(2)	0.0545	0.1628(6)	0.163216
	0.1605	0.3161(3)	0.472(1)	0.0447		
	0.1610	0.2823(4)	0.431(4)	0.0351		
	0.1615	0.2405(5)	0.426(6)	0.0255		
	0.1620	0.1903(8)	0.445(9)	0.0159		
Full, cooled	0.1250	0.5055(2)	0.5584(5)	0.157	0.1303(5)	0.13123
	0.1260	0.4507(2)	0.5165(10)	0.126		
	0.1270	0.3925(3)	0.4754(19)	0.094		
	0.1280	0.3273(5)	0.4352(40)	0.0635		
	0.1285	0.2898(7)	0.4160(62)	0.0483		
	0.1290	0.2470(10)	0.4000(99)	0.0327		
Quenched, uncooled	0.1600	0.326(4)	0.520(7)	0.0322	0.1617(1)	0.16191
	0.1610	0.206(4)	0.455(13)	0.0128		

Table 5.1: Meson masses for the configuration discussed in text.

Full correlation functions

Corr. function	cool	κ	m	w	b	χ^2
Pseudoscalar	0	0.1600	3.67577e-01	4.54506e+00	1.07018e+01	1.52363e-02
Vector	0	0.1600	5.11289e-01	7.37001e+00	1.35129e+01	6.12932e-03
Pseudoscalar	0	0.1610	3.04782e-01	4.57485e+00	1.24113e+01	1.73823e-02
Vector	0	0.1610	4.66675e-01	7.70165e+00	1.46338e+01	7.10922e-03
Pseudoscalar	20	0.1250	5.30744e-01	5.10255e+00	9.55370e+00	4.05560e-03
Vector	20	0.1250	5.99058e-01	7.72994e+00	1.26921e+01	1.08234e-03
Pseudoscalar	20	0.1285	3.14593e-01	4.87173e+00	1.25600e+01	9.31350e-03
Vector	20	0.1285	5.07063e-01	7.23912e+00	1.35451e+01	9.28434e-04
Pseudoscalar	20	0.1290	2.78628e-01	4.89639e+00	1.39631e+01	1.01811e-02
Vector	20	0.1290	5.29472e-01	7.94685e+00	1.45761e+01	1.16805e-03

Truncated correlation functions

Corr. function	cool	κ	N_v	m	w	χ^2
Pseudoscalar	0	0.1600	128	4.43973e-01	1.28681e+01	3.03466e-04
Pseudoscalar	0	0.1600	96	4.20370e-01	1.19193e+01	4.56920e-04
Vector	0	0.1600	128	6.94474e-01	3.69797e+00	1.23832e-03
Vector	0	0.1600	96	6.97106e-01	3.25622e+00	5.20885e-03
Pseudoscalar	0	0.1610	128	4.60775e-01	9.27689e+00	2.80168e-04
Pseudoscalar	0	0.1610	96	4.31284e-01	9.84512e+00	3.73627e-04
Vector	0	0.1610	128	6.63508e-01	4.34613e+00	1.42544e-03
Vector	0	0.1610	96	6.61804e-01	4.02443e+00	2.05692e-03
Pseudoscalar	0	0.1628	128	3.11332e-01	1.23271e+01	6.17862e-04
Pseudoscalar	0	0.1628	96	3.62355e-01	7.29528e+00	6.30676e-04
Vector	0	0.1628	128	6.02298e-01	1.68133e+01	1.05276e+00
Vector	0	0.1628	96	6.24801e-01	1.57323e+01	1.14086e+00
Pseudoscalar	20	0.1250	128	4.82490e-01	1.82229e+01	5.40265e-04
Pseudoscalar	20	0.1250	96	5.21136e-01	8.92229e+00	7.38465e-04
Vector	20	0.1250	128	7.98668e-01	1.77025e+00	6.15644e-03
Vector	20	0.1250	96	7.65802e-01	2.33539e+00	9.50609e-04
Pseudoscalar	20	0.1285	128	5.45822e-01	5.54051e+00	4.65273e-04
Pseudoscalar	20	0.1285	96	5.35643e-01	5.49999e+00	5.15762e-04
Vector	20	0.1285	128	6.97061e-01	2.85787e+00	9.44931e-04
Vector	20	0.1285	96	6.94689e-01	3.09413e+00	1.21252e-03
Pseudoscalar	20	0.1290	128	5.20004e-01	5.85261e+00	4.45830e-04
Pseudoscalar	20	0.1290	96	5.15167e-01	5.72019e+00	4.65942e-04
Vector	20	0.1290	128	6.94671e-01	2.89436e+00	8.96840e-04
Vector	20	0.1290	96	6.98770e-01	3.04270e+00	1.60514e-03
Pseudoscalar	20	0.1300	128	4.32990e-01	7.53830e+00	5.32361e-04
Pseudoscalar	20	0.1300	96	4.36831e-01	7.24053e+00	4.17893e-04
Vector	20	0.1300	128	7.87116e-01	1.65385e+00	8.80682e-03
Vector	20	0.1300	96	8.48759e-01	1.17100e+00	2.14198e-02

Table 5.2: Fit parameters for the correlation functions discussed in text. Parameters are defined in Eqs. (5.16) and (5.17), N_v denotes the number of eigenvectors used in propagator.

In Figures 5-18, 5-19 and 5-20 the full and truncated pseudo-scalar and vector correlation functions are presented for different configurations. The values of κ were chosen so that the pion masses are comparable in all cases. The full pseudo-scalar correlation function grows with distance, which is an indication of the strong attraction in pion channel. The full vector correlation function is almost constant up to the distance $r \approx 1$fm and falls off exponentially on larger distances. Unfortunately, the region $r \geq 1$fm is almost at the edge of the lattice and the results are not very reliable there. All data presented in these figures are obtained by the self-consistent solution described above.

The other curves on Figures 5-18, 5-19 and 5-20 correspond to the truncated correlation function computed with the different number of eigenvectors. In all cases, 16 eigenvectors are clearly not sufficient to describe the behaviour of correlation function even at intermediate distances. As the number of eigenvectors in the propagator increases, the truncated correlation function approaches the full correlation function at intermediate distances $\sim 1\text{fm} \approx 9a$ but the data beyond this point become very noisy. The maximum number of eigenvectors used, 128, gives a reasonable approximation to the ρ meson channel. In the pion channel the truncated correlation functions are higher then the full correlation function. Since the truncation of the propagator effectively introduces a new mass scale, the range of eigenvalues used for truncation, the truncated correlation function effectively corresponds to a full correlation function with a lighter mass. This effect is not important in vector channel but can be significant in the pseudo-scalar channel (see Section 5.7.3 for more discussion).

Figures 5-21, 5-22 and 5-23 show the dependence of the correlation functions on κ. This dependence is strong in the pion channel and weak in the ρ channel. Since κ is only a parameter in the truncated correlation function (see Eq. 5.13), those correlation functions can be computed with the value of $\kappa = \kappa_c$ ($m_q = 0$) without any critical slowing down. The truncated correlation function for $\kappa = \kappa_c$ is plotted on the same graphs but the numerical errors grow even more with $\kappa \to \kappa_c$ and the results are not reliable.

More definite statements on the behaviour of the truncated correlation functions require accumulation of additional statistics and more sophisticated treatment of image corrections or bigger lattices. On the other hand, even with the present data a reasonable qualitative approximation to the large distance behaviour of the correlation functions was obtained in the channels considered above. At most 128 eigenvectors were computed when the total number of eigenmodes was $4N_cV = 786432$, but the full and truncated correlation functions agree within 30%. This approximation shows that a significant amount of physical information is contained in the first few eigenmodes of the Dirac operator. The nature of those modes becomes clear after cooling, when the quantum noise is suppressed and only instantons survive. This result agrees with the previous calculations of correlation functions in uncooled and cooled lattices [17].

5.7.3 Comparison with Other Data

Since the correlation functions were studied in detail in the instanton liquid model [73, 60, 63] and were determined by phenomenological analysis of experimental data, the full and truncated correlation functions calculated above can be compared with those data.

In Figure 5-24 the results of this work are compared with the instanton liquid model results and phenomenological fits to experimental data from [73] for the channels considered in this work. The other two curves were taken from [17]. In that paper the correlation functions were calculated using the quenched approximation with $\beta = 5.7$ on a $16^3 * 32$ lattice. This choice of parameters leads to the lattice spacing (as determined by the nucleon mass) $a_N = 0.165$fm.

Because of the strong dependence of correlation function on the quark mass in the pseudo-scalar channel, the correlation function with approximately the same quark mass ($m_q \approx 25$MeV) from ref. [17] was used for comparison. The dependence of the truncated correlation function on the quark mass is weaker because the truncation effectively sets the mass scale in the correlation function. The truncated correlation function agrees with both phenomenological data and the correlation function extrapolated to light quark masses.

In the vector channel the dependence on the quark mass is weak and the full correlation function is close to the extrapolated one even for large quark masses. The truncated correlation function only partially reproduces the full correlation function even at large distances ($\approx$ 1fm) which may indicate that 128 eigenvectors is not enough for this configuration. The agreement of the truncated correlation function, the full correlation function and the phenomenological data is better for full QCD configurations. One reason for this behaviour may be the suppression of instantons by dynamical quarks. Although there are comparable numbers of instantons in Figure 4-9 for quenched and unquenched configurations, the eigenvalues are concentrated closer to the real axis in the quenched case. Hence the zero-mode zone consists of fewer modes (less then 128) for the unquenched case.

On the graphs for both channels the curve labeled "resonance" comes from the fit to the spectral representation of the correlation function [17]. The spectrum was fitted with the contribution from the single isolated resonance (corresponding to the lowest-mass particle in the given channel) and the suitably parameterized continuum contribution. Then the parameters of the resonance were approximated to continuum limit. The contribution of this single resonance dominates the correlation function in the same region where the truncated correlation function approaches the full correlation function. This coincidence suggests that the lowest-mass mesons (π and ρ) are described by instanton dynamics and are excitations of the chiral condensate. The heavier particles must involve additional degrees of freedom which are outside the zero-mode zone. (I am grateful to Edward Shuryak for clarification of this point).

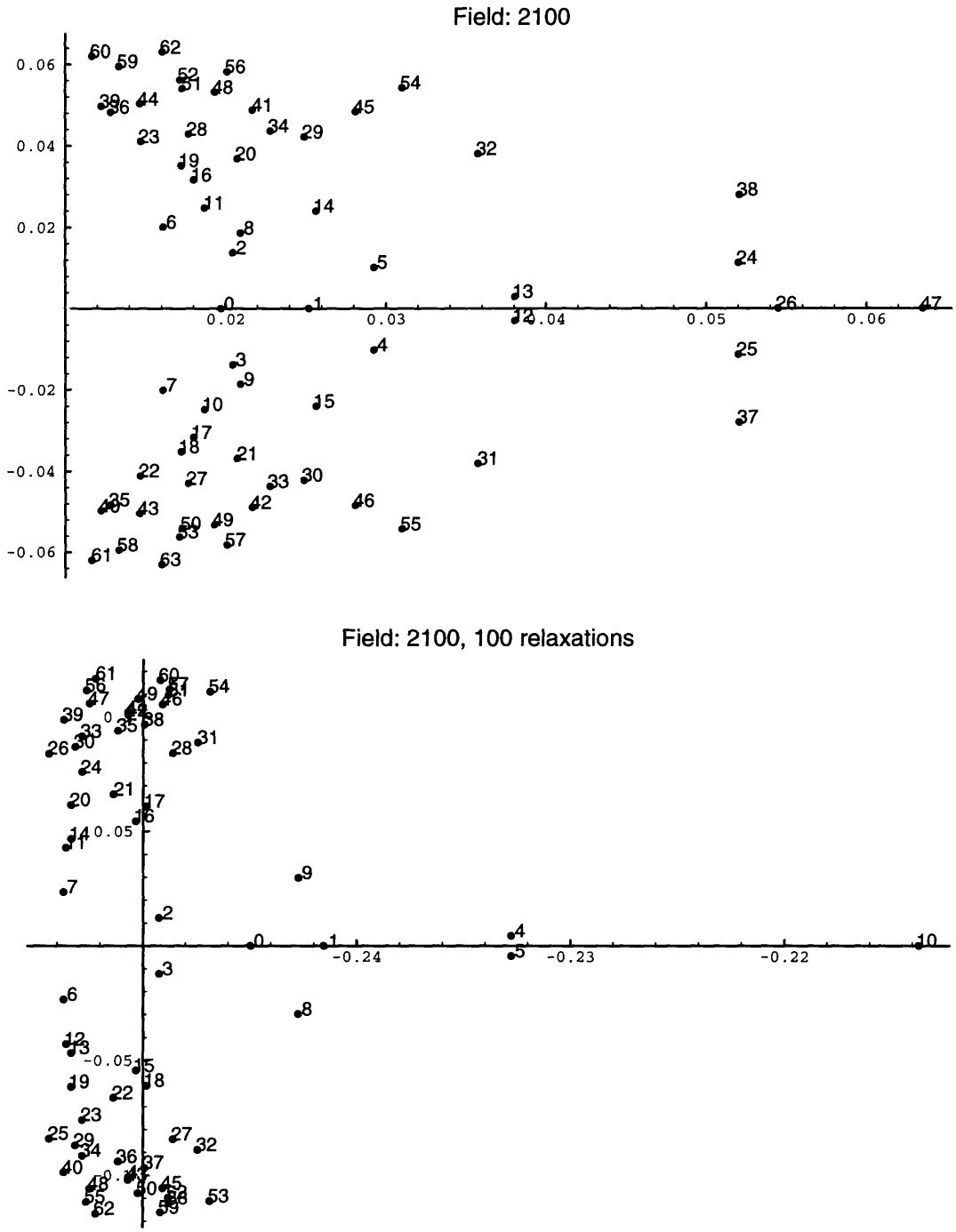

Figure 5-8: The lowest 64 modes of the Dirac operator on one selected lattice after 0 and 100 relaxation steps. In both cases, $\kappa = 0.1600$ and the large negative values on the lower graph indicate that κ_c for this configuration is much lower (≈ 0.125).

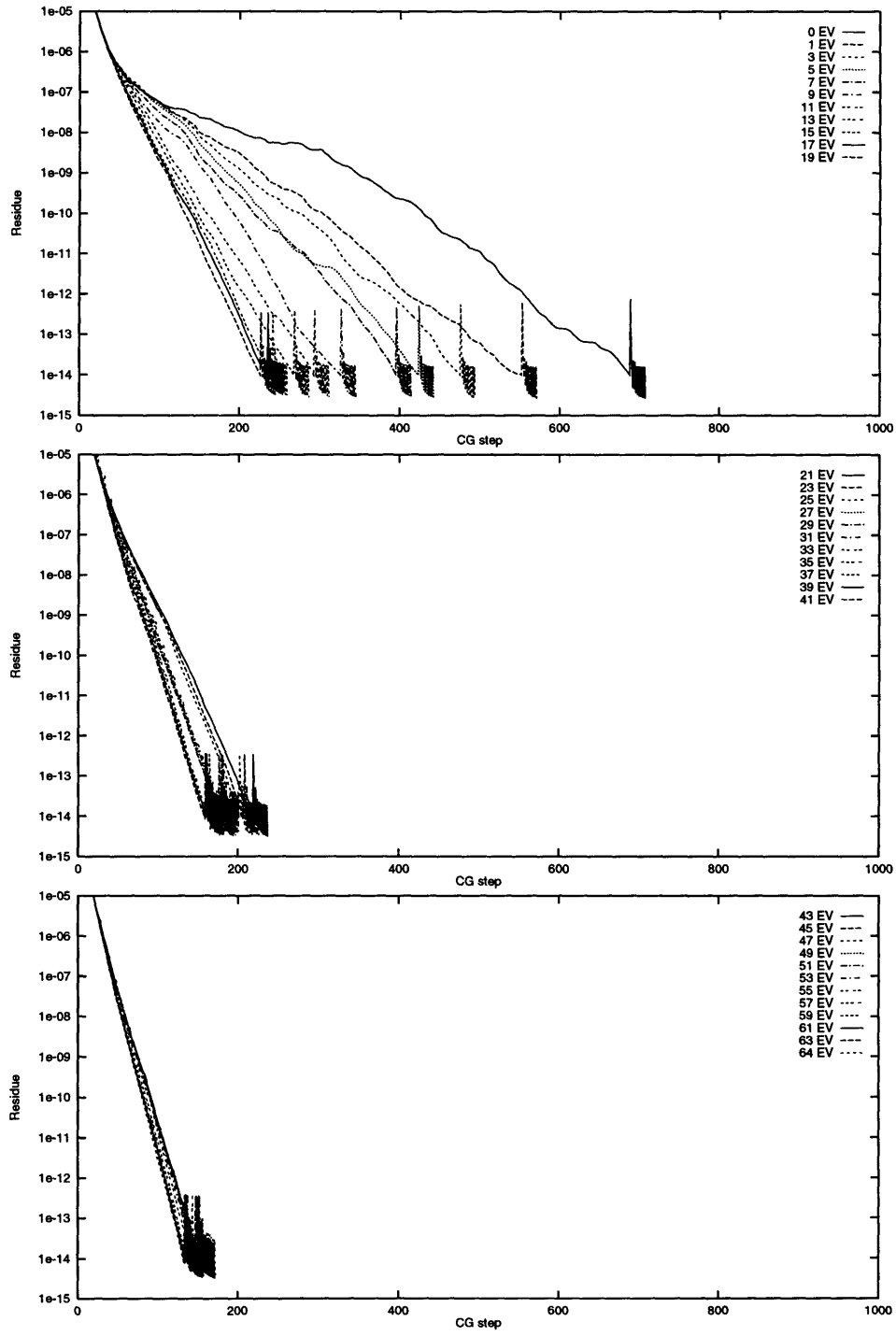

Figure 5-9: Conjugate gradient convergence for $\kappa = 0.1615$ and different numbers of eigenmodes subtracted.

In each case, the number of subtracted eigenmodes is shown in the key and the residue, $r = \|D\chi - \psi\|^2/\|\chi\|^2$ is plotted as a function of the number of CG steps.

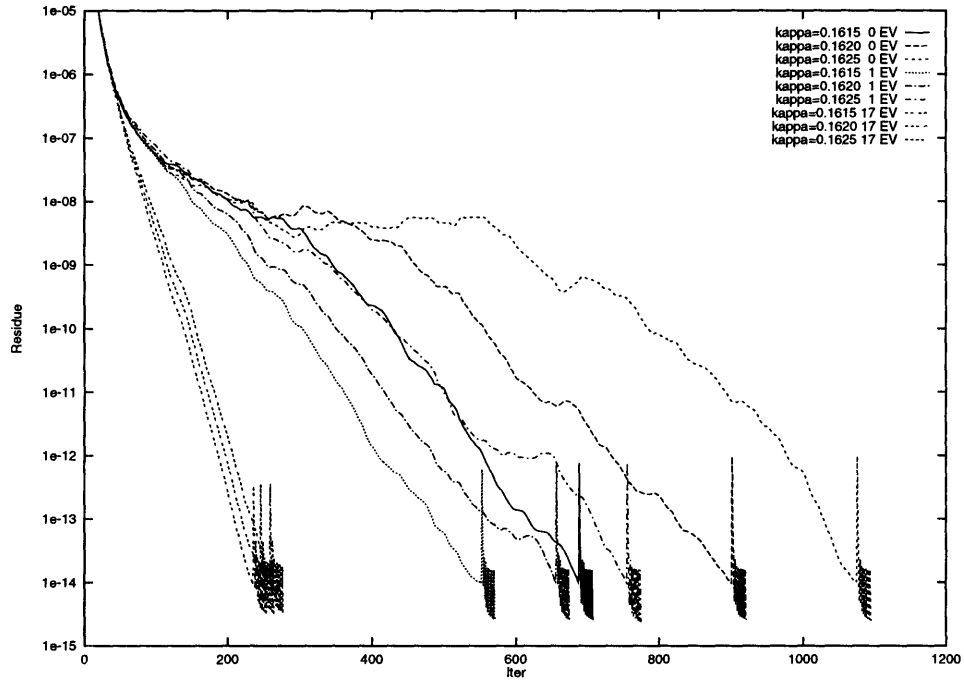

Figure 5-10: Convergence of CG as a function of iteration with 0,1 and 17 eigenvectors subtracted for different values of κ.

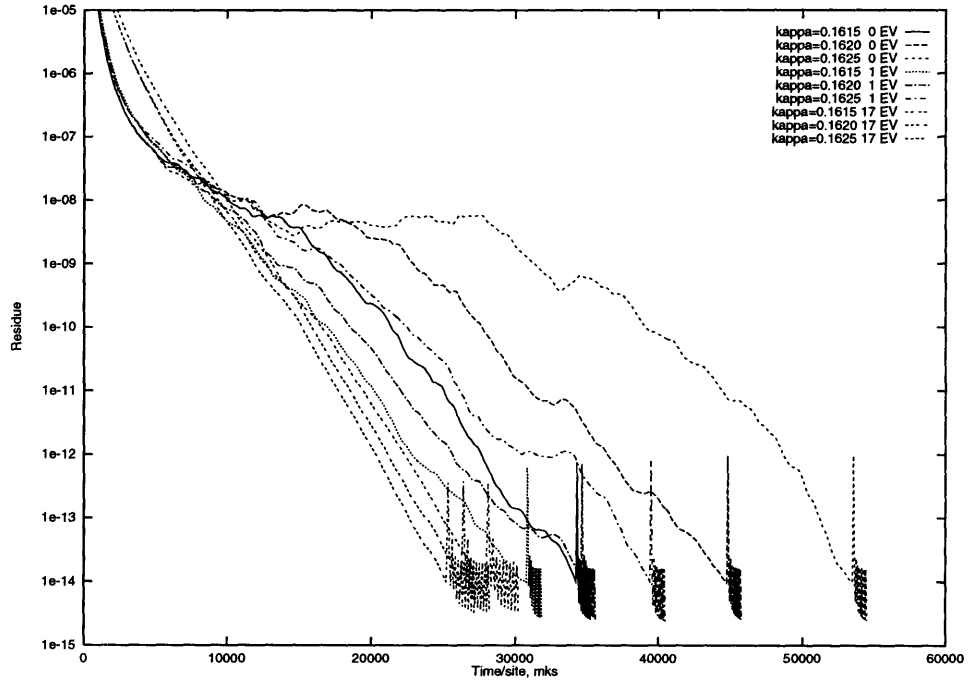

Figure 5-11: Convergence of CG as a function of computation time with 0,1 and 17 eigenvectors subtracted for different values of κ.

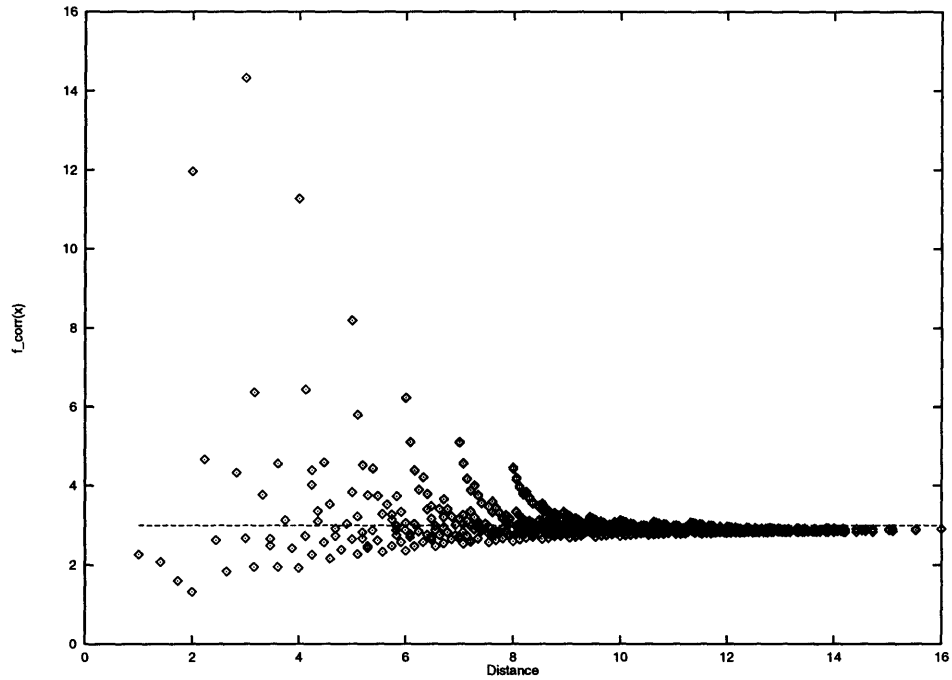

Figure 5-12: Lattice anisotropy correction function. In the limit in which the lattice spacing $a \to 0$, the function would be a constant $N_c = 3$.

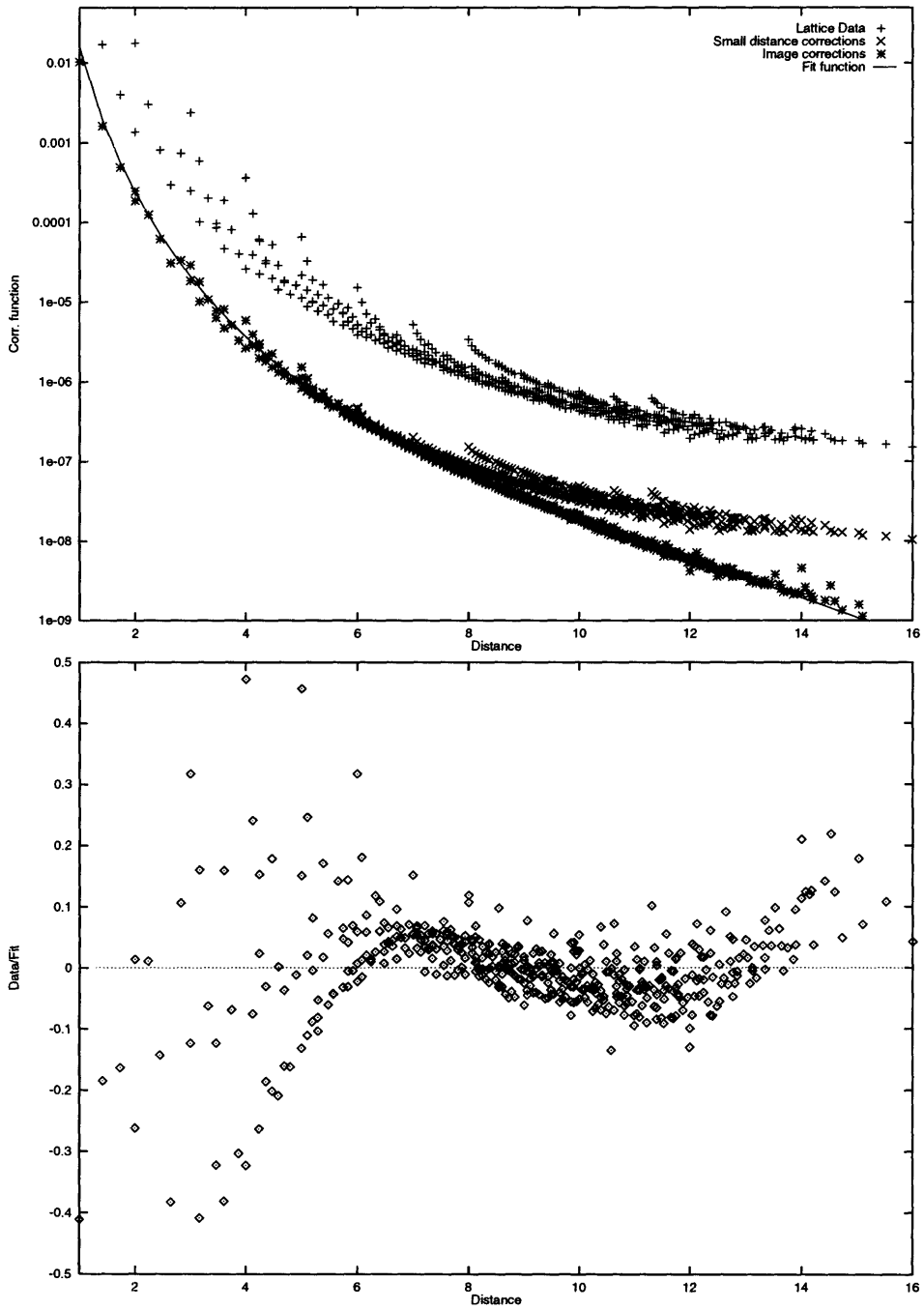

Figure 5-13: The fit to the full vector current correlation function for $\kappa = 0.1610$. The upper figure shows the raw data for the lattice correlation function averaged over equivalent points (pluses), the same data after application of the small-distance anisotropy corrections (crosses), the image-corrected data (stars) and the final form of the fit function (solid line). The corrected data are shifted down by an arbitrary amount for clarity. The bottom graph shows the log of the ratio of fit points to actual data.

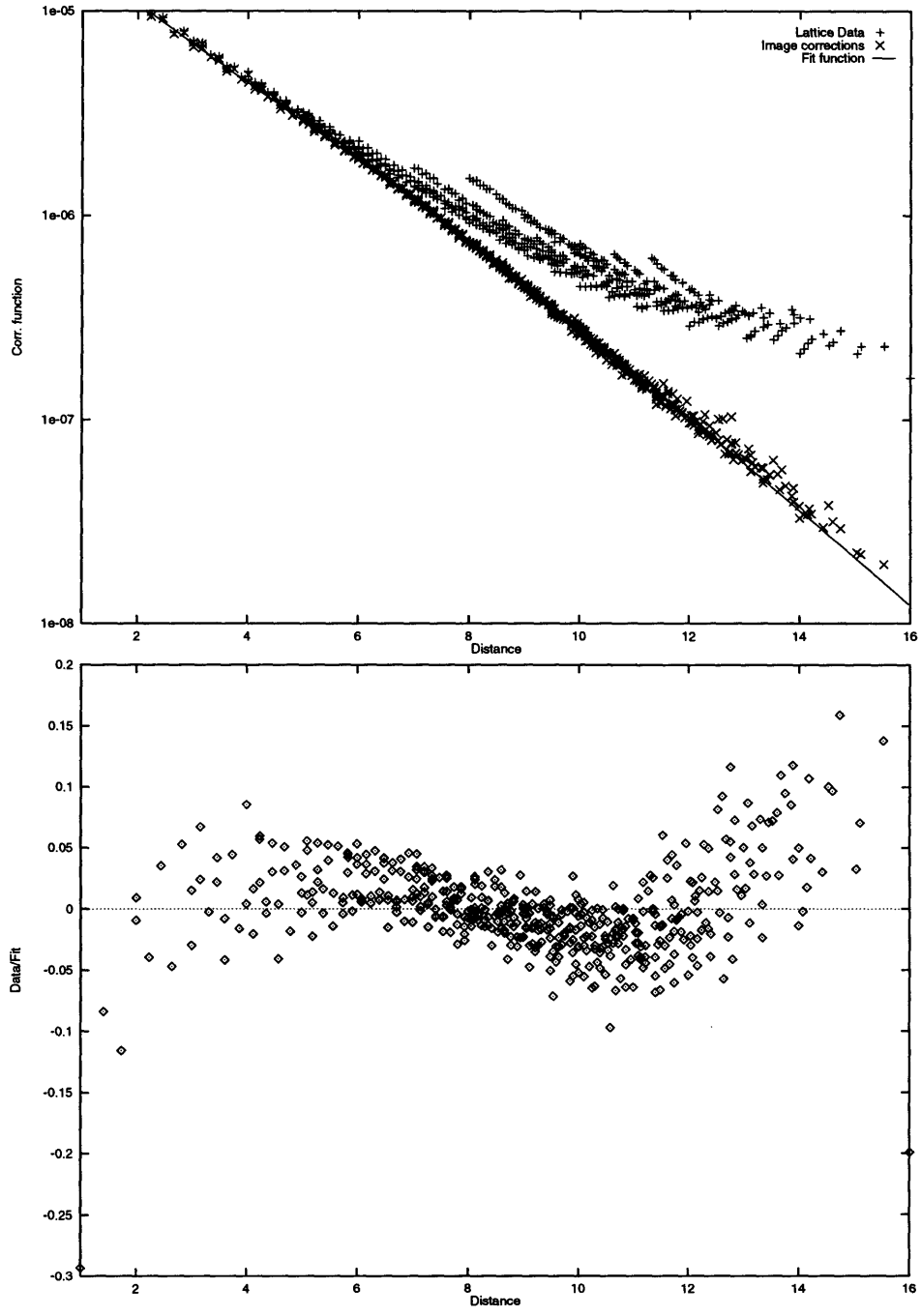

Figure 5-14: The fit to the truncated vector current correlation function for $\kappa = 0.1610$ and 128 eigenvectors. All curves and symbols are as defined in Figure 5-13.

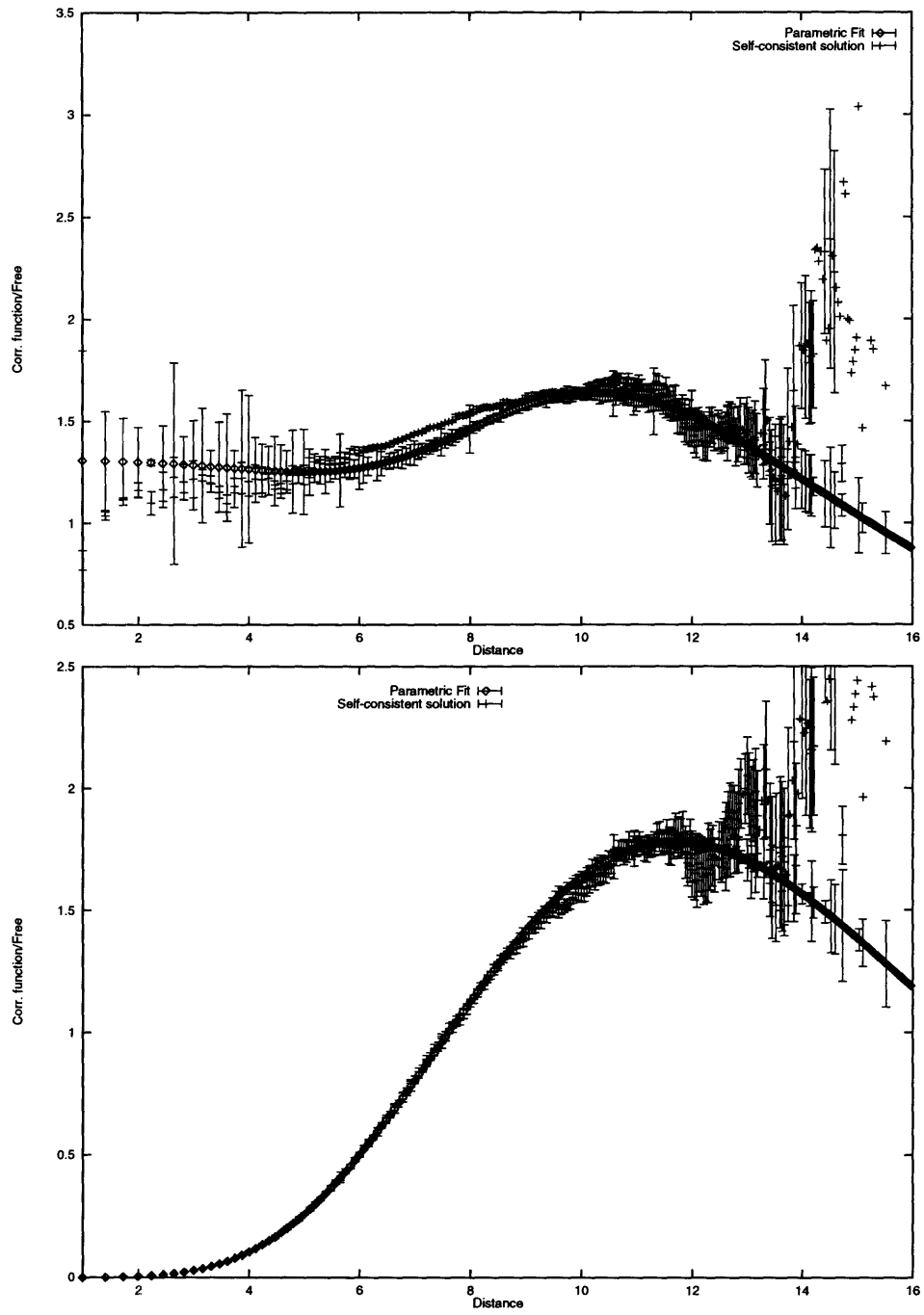

Figure 5-15: Comparison of parametric fit with self-consistent solution (Eq. 5.18)
The upper Figure shows the normalized full correlation function (vector channel, $\kappa = 0.1600$) calculated by solving Eq. (5.18) and the parametric fit to the same correlation function. The two approximations agree within errors which were estimated by the spread of points. The lower graph shows the same correlation function calculated with 128 eigenvectors.

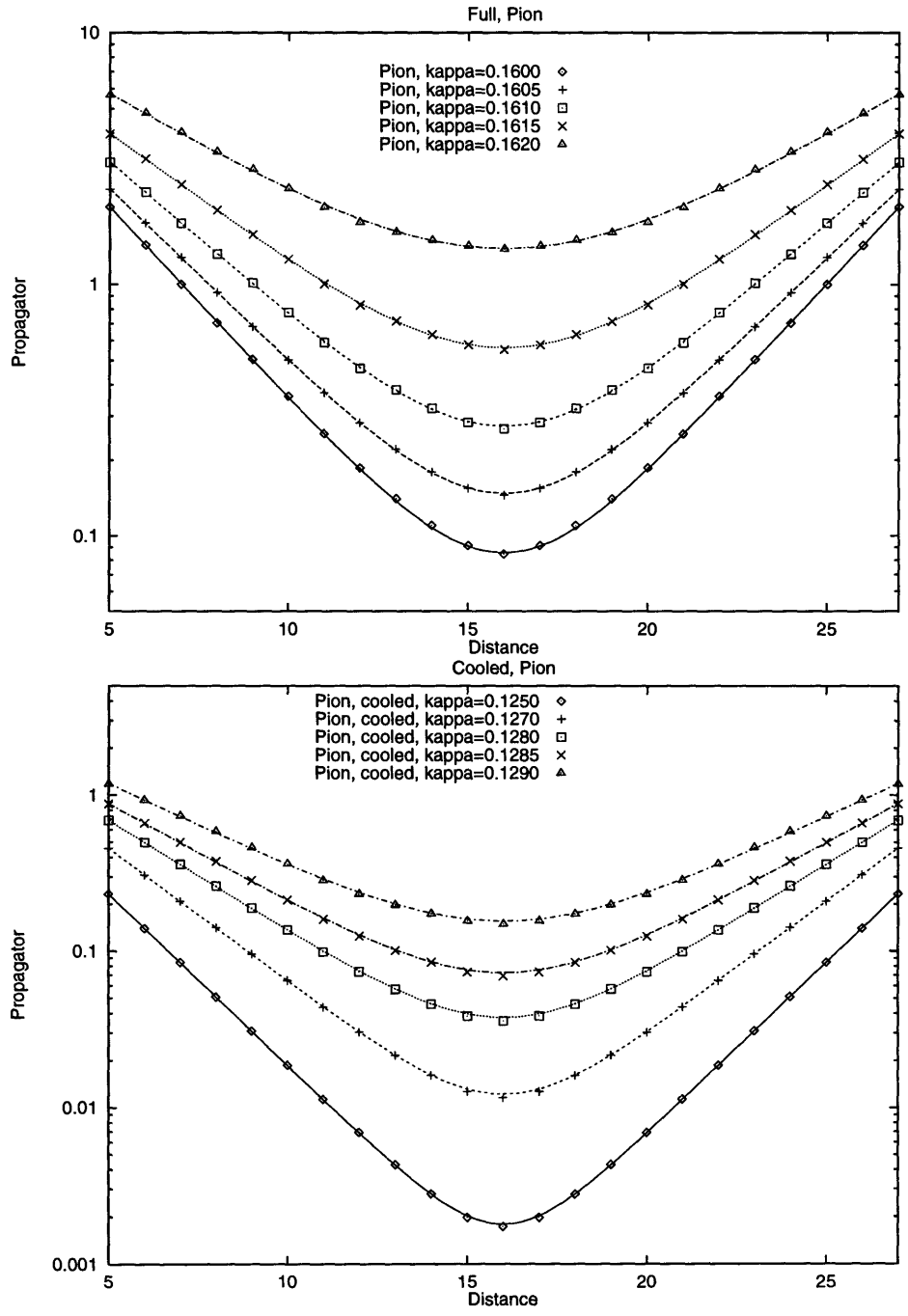

Figure 5-16: Fit of the zero-momentum projection of the pion propagators to the sum of exponentials to determine masses for uncooled and cooled configurations

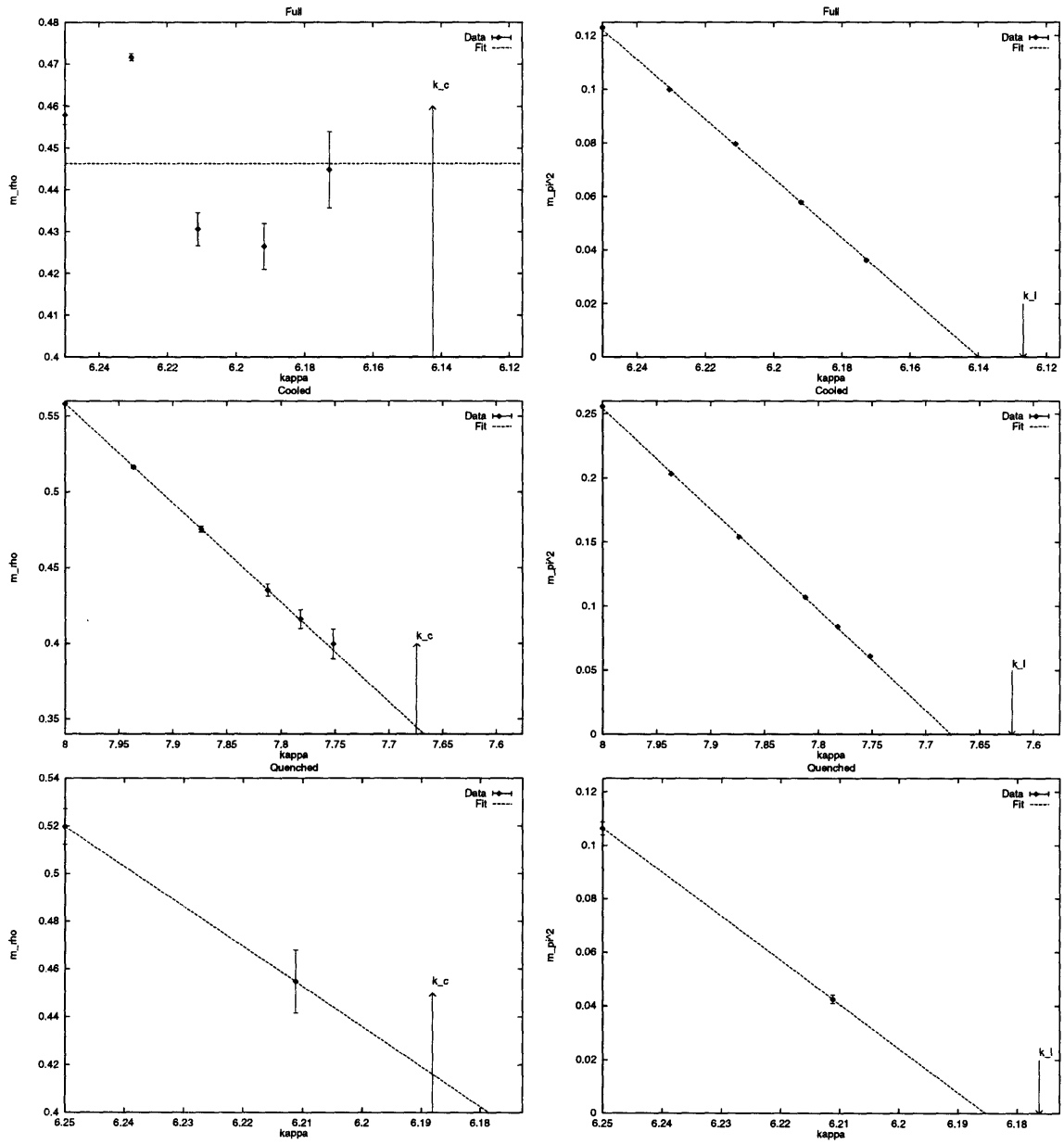

Figure 5-17: Fit of m_π^2 vs. κ to determine κ_c for uncooled and cooled configurations and estimate of m_ρ.

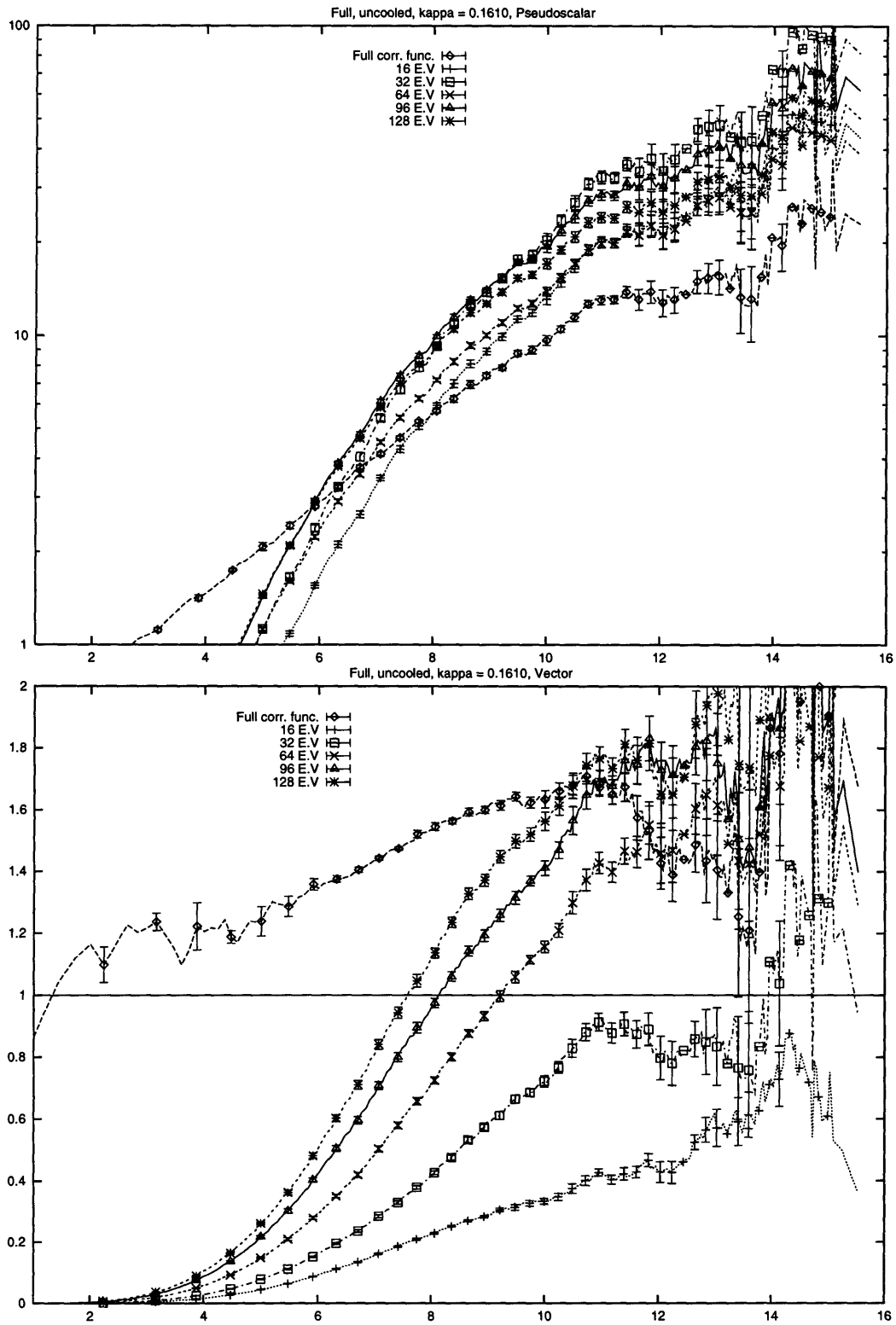

Figure 5-18: Normalized correlation functions for a full QCD configuration, $\kappa_v = 0.1610$, $m_\pi a = 0.2823(4)$, $m_q a = 0.0351$

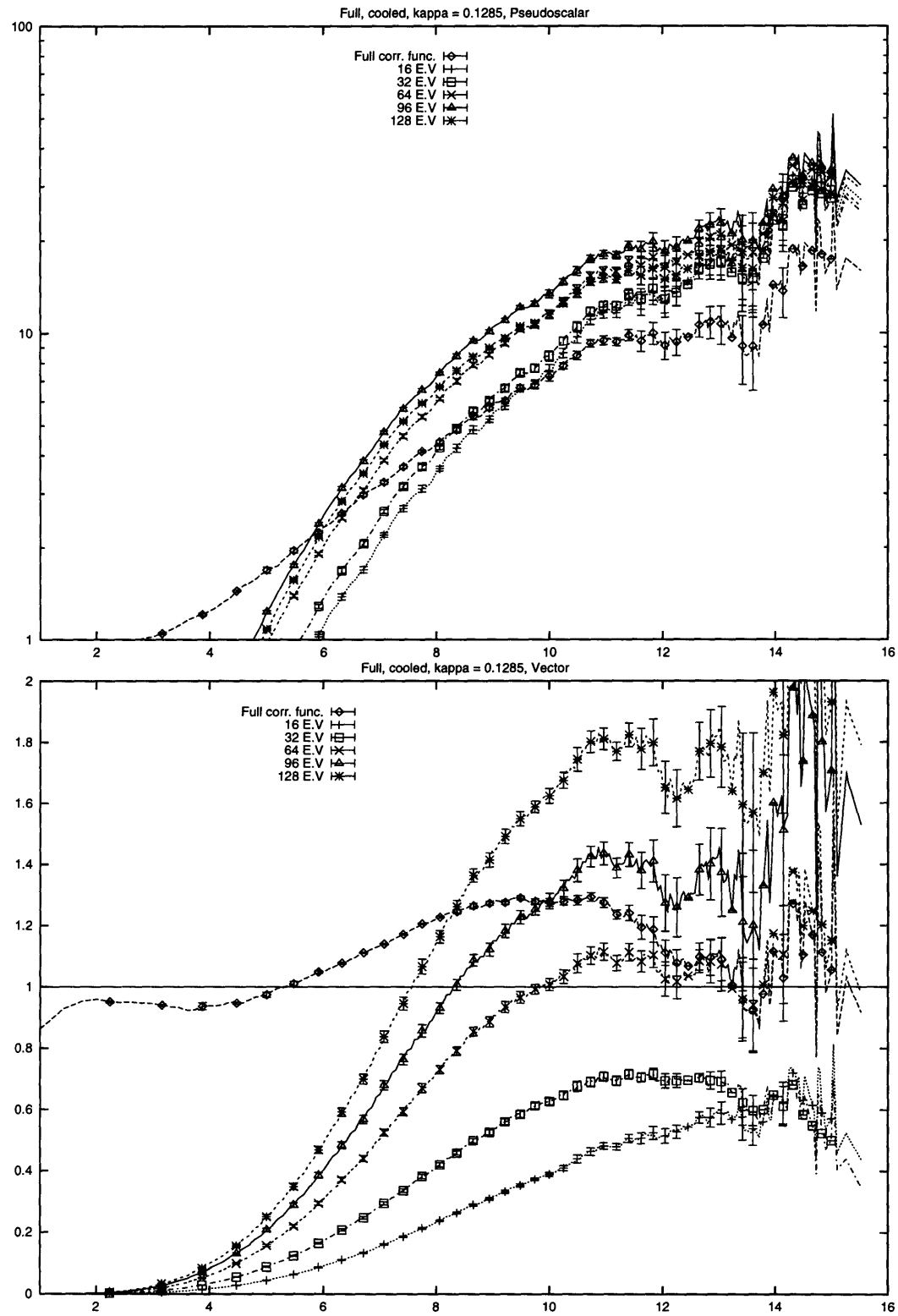

Figure 5-19: Normalized correlation functions for a full QCD configuration, relaxed 20 steps with $\epsilon = 0.025$, $\kappa_v = 0.1285$, $m_\pi a = 0.2898(7)$, $m_q a = 0.0483$

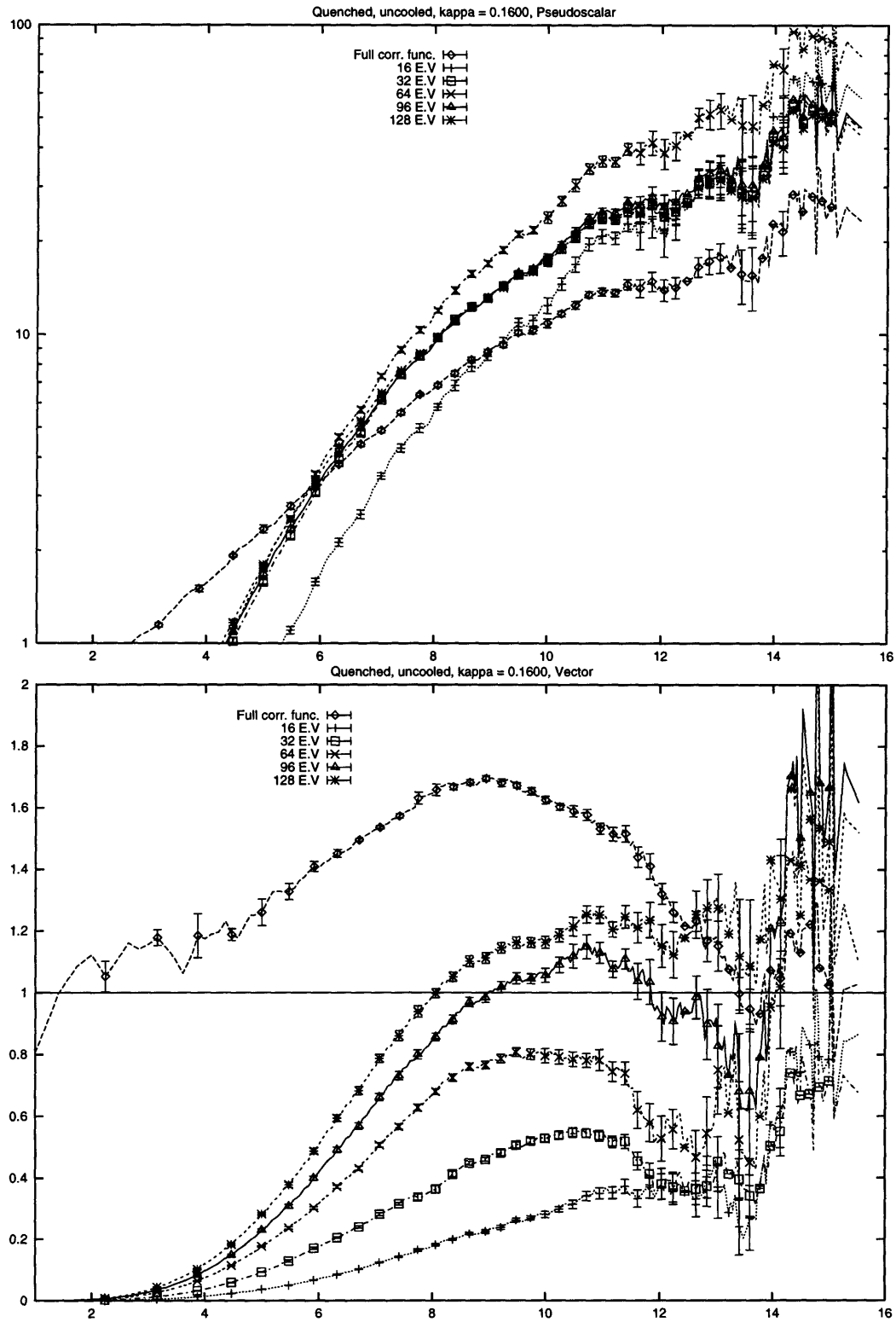

Figure 5-20: Normalized correlation functions for a quenched configuration, $\kappa_v = 0.1600$, $m_\pi a = 0.326(4)$, $m_q a = 0.0322$

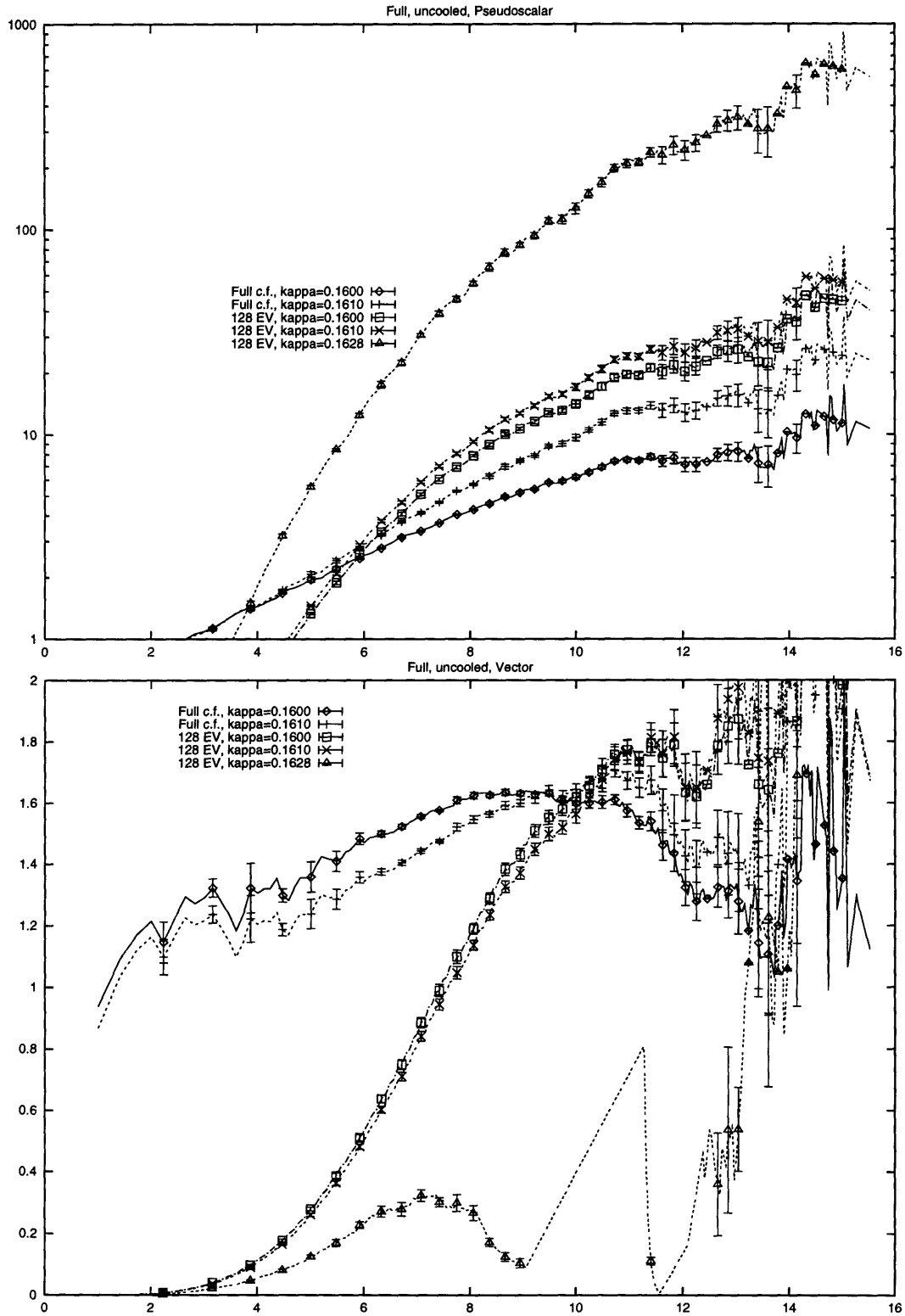

Figure 5-21: Comparison of full and truncated correlation functions for different κ for a full QCD configuration. The value of $\kappa = 0.1628 \approx \kappa_c$ corresponds to a very light quark mass $m_q \leq 20$MeV.

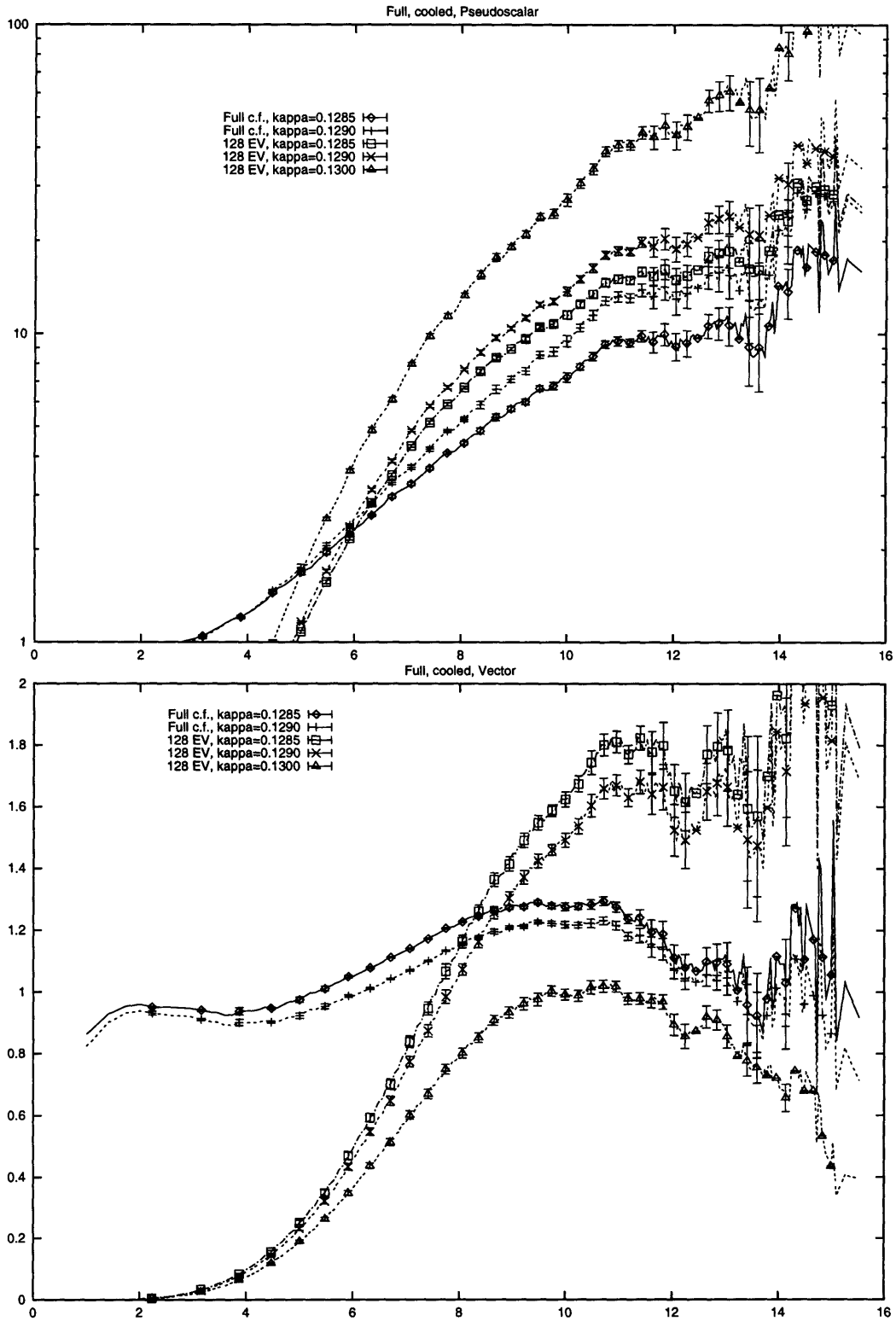

Figure 5-22: Comparison of full and truncated correlation functions for different κ for a cooled configuration. The value of $\kappa = 0.1300 \approx \kappa_c$ corresponds to a very light quark mass $m_q \leq 42$MeV.

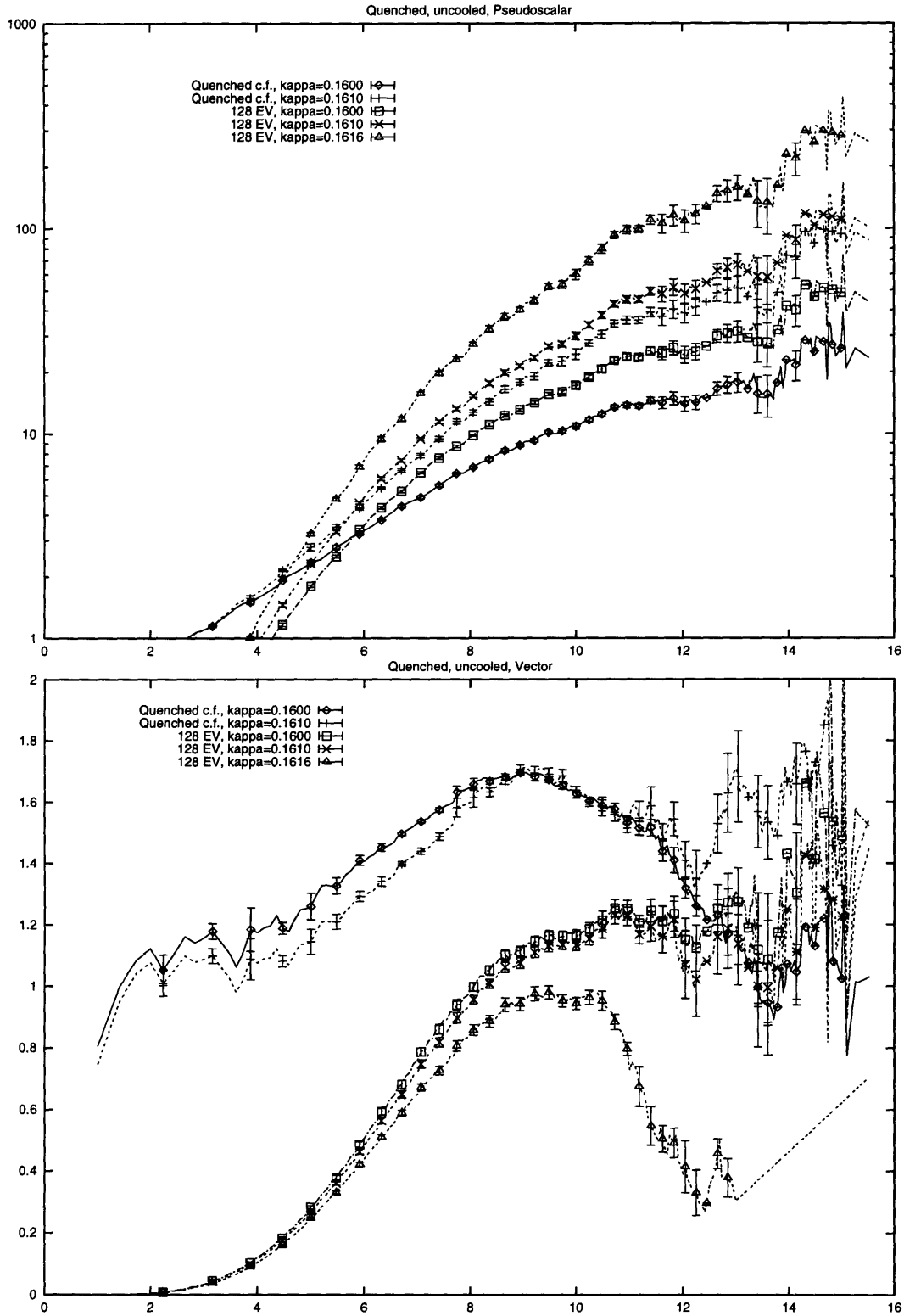

Figure 5-23: Comparison of full and truncated correlation functions for different κ for a quenched configuration. The value of $\kappa = 0.1616 \approx \kappa_c$ corresponds to a very light quark mass $m_q \leq 7$MeV.

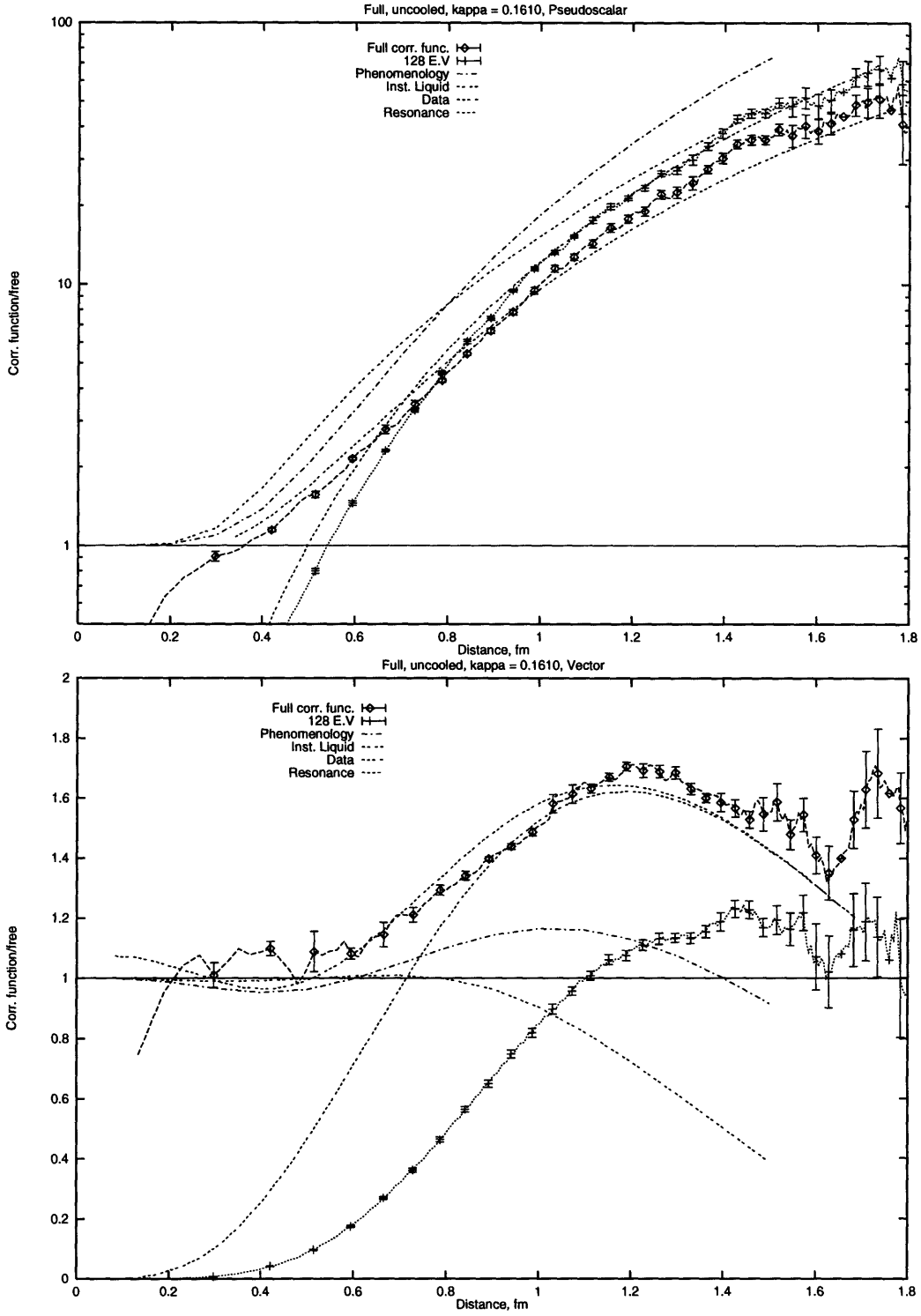

Figure 5-24: Comparison of the full and truncated correlation functions with phenomenological results and the instanton liquid model. The full and truncated correlation function correspond to a quenched configuration with $\kappa = 0.1610$. The phenomenological fit and instanton liquid model result were taken from [73]. The curve labeled "Data" was determined in [17] from am ensemble of quenched configurations and the curve labeled "Resonance" is the contribution to the correlation function from the lowest resonance. All correlation functions are plotted as a function of the physical distance (in Fermi).

Chapter 6

Summary and Outlook

This work has studied the effects of long-range excitations in QCD. The instanton distribution in QCD was determined with and without internal fermions (Chap. 4) and no statistically significant difference between the two distributions was found. It is possible that this similarity resulted from the large sea quark mass we had to use ($m_q \approx 100$MeV) and future study of larger lattices which allow lighter quarks is needed. The instanton distribution changes significantly under relaxation process which complicates the exact determination of the instanton distribution. The finite-volume effects dominate the dynamics of well-separated instantons under relaxation but are not significant for the small amount of relaxation needed to study the dense instanton liquid.

The rôle of the low-eigenvalue modes in the Hybrid Monte-Carlo dynamics (Chap. 3) has been studied and the efficiency of this method was increased significantly by using few recent vectors to approximate the solution of the inverse of Dirac operator. The algorithm to find the low-eigenvalue modes of the Dirac operator was developed and tested on different configurations. This algorithm was used to study the spectrum of the Dirac–Wilson operator on the lattice (Chap. 5). Restricting the Dirac operator to the complement of few lowest modes was shown to completely eliminate critical slowing down in the Conjugate Gradient method (Sec. 5.6). Since the eigenvectors do not depend on the quark mass (or hopping parameter κ), the eigenvectors computed with one value of κ can be used to speed up the CG convergence for many sources and κ.

The calculated low-λ eigenmodes of the Dirac operator were used to approximate the quark propagator by truncating the spectral representation. The point–to–point correlation function obtained using the truncated propagator were compared with correlation functions calculated by the traditional CG method. The full and truncated correlation functions agree qualitatively at large distances although the number of eigenvectors used is many orders of magnitude smaller than the total number of modes in the spectrum. This result shows that the essential part of QCD dynamics on intermediate distance $r \approx 1$fm is described by the low-eigenvalue part of the spectrum. The

comparison of the correlation functions in uncooled and cooled configuration shows that quark propagation can be described by instanton zero modes. The truncated correlation functions approximate the contribution to the correlation function from the peak in the spectrum corresponding to the lowest–mass particle in the particular channel. Taken together those results strongly suggest that the lowest mass particles in different meson channels can be described by the dynamics of instantons in the QCD vacuum.

Several questions that were not addressed in this work merit further study. The instanton distribution could be better determined by using the streamline solutions to fit the action and topological charge densities on the lattice. The exact form of this solution depends on the relative orientations of instantons which were left undetermined in the approximation used in Chapter 4. Alternatively, better variables could be found to replace the action and topological charge densities for the purpose of fitting the instanton background. Knowing the instanton orientations would permit a better study of eigenvalue splitting (see Chapter 5) although the benefits of such a study are not clear. The developed algorithm to determine the low-λ part of the spectrum opens interesting opportunities for studying the lattice QCD configurations by fermionic methods. Since the chiral instanton zero modes correspond to instantons carrying topological charge, the topological charge of the lattice can be determined by counting the zero modes. Such a definition could be compared with other methods of defining the topological charge of the lattice. Other physical quantities like the chiral condensate, baryonic correlation functions and particle structure functions can also be studied using the truncation of the spectrum.

Appendix A

Conventions and Useful Formula

In this Appendix relevant definitions and formula are collected for completeness.

Sigma matrices are very standard:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \sigma_4 = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} \tag{A.1}$$

This work uses the chiral representation of gamma matrices. Any other representation can be obtained from those by the appropriate rotation in the spin space. All gamma matrices $\gamma_{1...5}$ are Hermitian and, of course, anti-commute with each other.

$$\gamma_1 = \begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{A.2}$$

$$\gamma_2 = \begin{pmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \tag{A.3}$$

$$\gamma_3 = \begin{pmatrix} 0 & \sigma_3 \\ \sigma_3 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \tag{A.4}$$

$$\gamma_4 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & i \\ -i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix} \tag{A.5}$$

$$\gamma_5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{A.6}$$

The t'Hooft symbols [79] enter the definition of the instanton field and mix spinor and space indices. In fact, they are the coefficients of the decomposition $O(4) \approx SU(2) \times SU(2)$.

$$\eta^a_{\mu\nu} = \begin{cases} \epsilon^{a\mu\nu} & \mu,\nu \neq 4 \\ \delta^{a\mu} & \mu \neq 4, \nu = 4 \\ -\delta^{a\nu} & \mu = 4, \nu \neq 4 \end{cases} \tag{A.7}$$

$$\bar{\eta}^a_{\mu\nu} = \begin{cases} \epsilon^{a\mu\nu} & \mu,\nu \neq 4 \\ -\delta^{a\mu} & \mu \neq 4, \nu = 4 \\ \delta^{a\nu} & \mu = 4, \nu \neq 4 \end{cases} \tag{A.8}$$

Those symbols have the following properties [73, 79]

$$\eta^a_{\mu\nu}\eta^b_{\mu\nu} = 4\delta^{ab} \tag{A.9}$$

$$\eta^a_{\mu\nu}\eta^a_{\mu\rho} = 3\delta_{\nu\rho} \tag{A.10}$$

$$\eta^a_{\mu\nu}\bar{\eta}^b_{\mu\nu} = 0 \tag{A.11}$$

$$\eta^a_{\mu\nu}\eta^b_{\mu\sigma} = \delta^{ab}\delta_{\nu\sigma} + \epsilon^{abc}\eta^c_{\nu\sigma} \tag{A.12}$$

$$\eta^a_{\mu\nu}\eta^a_{\rho\sigma} = \delta_{\mu\rho}\delta_{\nu\sigma} - \delta_{\mu\sigma}\delta_{\nu\rho} + \epsilon_{\mu\nu\rho\sigma} \tag{A.13}$$

$$\bar{\eta}^a_{\mu\nu}\bar{\eta}^a_{\rho\sigma} = \delta_{\mu\rho}\delta_{\nu\sigma} - \delta_{\mu\sigma}\delta_{\nu\rho} - \epsilon_{\mu\nu\rho\sigma} \tag{A.14}$$

$$\epsilon_{abc}\eta^b_{\mu\nu}\eta^c_{\rho\lambda} = \delta_{\mu\rho}\eta^a_{\nu\lambda} - \delta_{\mu\lambda}\eta^a_{\nu\rho} + \delta_{\nu\lambda}\eta^a_{\mu\rho} - \delta_{\nu\rho}\eta^a_{\mu\lambda} \tag{A.15}$$

$$\epsilon_{\lambda\mu\nu\sigma}\eta^a_{\rho\sigma} = \delta_{\rho\lambda}\eta^a_{\mu\nu} + \delta_{\rho\nu}\eta^a_{\lambda\mu} + \delta_{\rho\mu}\eta^a_{\nu\lambda} \tag{A.16}$$

Appendix B

Matrix Algorithms

This Appendix describes in detail the algorithms used throughout this thesis. Since conjugate gradient (CG), biconjugate gradient (bicg), minimal residue (MR) and Lanczos algorithms were extensively used in lattice calculations [33, 34, 55, 58], they are not described here.

B.1 Even–Odd Pre-conditioning of the Dirac–Wilson Operator

First, divide all lattice sites into two sets by the "parity"

$$P(x) = (x + y + z + t) \bmod 2, \tag{B.1}$$

all sites with $P(x) = 0$ are called "even" and all sites with $P(x) = 1$ are called "odd". The Dirac-Wilson operator (Eq. 2.6) has two pieces, the diagonal piece which was normalized to 1 and the off-diagonal piece which connects even and odd sites. In matrix notation it can be represented as

$$D\psi = \begin{pmatrix} 1 & -\kappa K_{eo} \\ -\kappa K_{oe} & 1 \end{pmatrix} \begin{pmatrix} \psi_e \\ \psi_o \end{pmatrix} \tag{B.2}$$

The determinant of D in a QCD partition function (Eq. 2.10) can be simplified to the determinant of even-to-even operator

$$M = 1 - \kappa^2 K_{eo} K_{oe} \tag{B.3}$$

We shall see that the same operator enters into all computations in the theory.

The inversion problem $D\chi = \psi$ written in the matrix notation is

$$\begin{cases} \chi_e - \kappa K_{eo}\chi_o = \psi_e \\ -\kappa K_{oe}\chi_e + \chi_o = \psi_o \end{cases} \tag{B.4}$$

The second equation can be solved for χ_o giving

$$\chi_o = \psi_o + \kappa K_{oe}\chi_e \tag{B.5}$$

which leads to the inversion problem of M with modified source:

$$\left(1 - \kappa^2 K_{eo}K_{oe}\right)\chi_e = \psi_e + \kappa K_{eo}\psi_o \tag{B.6}$$

The eigenvalue problem for D can be tackled in the same way:

$$D\psi = \lambda_D\psi \tag{B.7}$$

$$\begin{cases} \psi_e - \kappa K_{eo}\psi_o = \lambda_D\psi_e \\ -\kappa K_{oe}\psi_e + \psi_o = \lambda_D\psi_o \end{cases} \tag{B.8}$$

$$\psi_o = \frac{1}{1-\lambda_D}\kappa K_{oe}\psi_e \tag{B.9}$$

$$\psi_e - \frac{\kappa^2}{1-\lambda_D}K_{eo}K_{oe}\psi_e = \lambda_D\psi_e \tag{B.10}$$

$$M\psi_e = \left(1 - \kappa^2 K_{eo}K_{oe}\right)\psi_e = \lambda_M\psi_e \tag{B.11}$$

(B.12)

We also get the important equation relating eigenvalues of M and D

$$1 - \lambda_M = (1 - \lambda_D)^2 \tag{B.13}$$

Note that each eigenvalue of M corresponds to 2 distinct eigenvalues of D. If one of those eigenvalues is close to zero, the other is close to 2 and is not interesting in the physical region. When the spectrum $\lambda_i(\kappa)$ is known for one value of κ, the spectrum for all other values can be obtained by simple rescaling

$$\lambda_i(\kappa') = \frac{\kappa'}{\kappa}\lambda_i(\kappa) + \frac{\kappa - \kappa'}{\kappa} \tag{B.14}$$

B.2 Implementation of the k-step Arnoldi method to find the lowest eigenvectors of the Dirac operator

Since the Wilson-Dirac operator is not Hermitian, the simple Lanczos methods [58, 20]. can not be applied to it. One option is to use the combined Hermitian operators $M_5 = \gamma_5 M$ or, equivalently, $M^\dagger M = M_5^2$. The operator M_5 is simpler but not positive-definite and extra care should be taken in the algorithms using it. In this standard way up to all eigenvalues can be found [49].

Unfortunately, since the operators γ_5 and M do not commute, the spectrum of M_5 has no direct relation to the spectrum of D and only in the simultaneous limits $a \to 0$ and $m \to 0$ the two spectra coincide. Even worse, the spectrum of M_5 has non-trivial dependence on κ and knowing the spectrum at one value of κ does not help in determining the spectrum at other values of κ.

As discussed in Chapter 5, the few lowest eigenmodes of the Wilson-Dirac operator have particularly significant impact on physics, so it is essential to be able to calculate them. The lowest eigenvalues of M_5 were studied by the modified conjugate-gradient method [47, 48, 50].

The k-step Arnoldi method developed by D. Sorensen [76] was applied to find the lowest eigenvalues of the true Wilson-Dirac operator. Since it was not discussed previously in the lattice physics literature, the detailed description of it is given below. One of the important advantages of this algorithm is almost complete insensitivity to κ: once the eigenvalues for one value of κ are known, the spectrum for any other κ can be determined. Using this feature one does not need to know the critical κ while computing eigenvalues (see Eq. B.14).

First, starting from the arbitrary initial vector v_1 build the basis up to v_k by applying the operator M and orthogonalizing the result to all previous vectors. If M were Hermitian, this iteration would lead to a tri-diagonal matrix as in Lanczos algorithm. In the general case, the matrix of coefficients is upper-Hessian [58]. In vector components the process is defined by

$$\begin{aligned} Mv_1 &= v_1 h_{11} + v_2 h_{21} \\ Mv_2 &= v_1 h_{12} + v_2 h_{22} + v_3 h_{32} \\ &\dots \\ Mv_k &= v_1 h_{1k} + \dots + v_k h_{kk} + r_k \end{aligned}$$

where r_k is a residual vector. In the algorithmic form the process is defined by the sequence of steps

$$h_{k+1,k} \leftarrow |r_k| \tag{B.15}$$

$$v_{k+1} \leftarrow r_k/|r_k| \tag{B.16}$$

$$h_{i,k+1} \leftarrow \langle v_i|M|v_{k+1}\rangle \tag{B.17}$$

$$r_{k+1} \leftarrow Mv_{k+1} - \sum_i v_i h_{i,k+1} \tag{B.18}$$

Denoting the matrix of vectors $V^{(k)} = (v_1, \ldots, v_k)$, the residual vector $R^{(k)} = (0, \ldots, r_k)$ the relations above can be written in a matrix form

$$MV = VH + R \tag{B.19}$$

If this algorithm is continued to $k = \dim(M)$ the residue would be zero and eigenvalues of M are eigenvalues of H. To finish the calculations this way *all* vectors v_k must be stored. The full spectrum of Dirac–Wilson operator has been found by a similar procedure [61]. In the k-step Arnoldi algorithm the residue r is driven to zero by successive iterations and vectors v_i go toward the span of the few eigenvectors.

The matrix H can be decomposed into a product of a unitary and right-triangular matrix U for any number μ: $H - \mu 1 = QU$ [58]. The matrix $H' = UQ + \mu 1 = Q^\dagger HQ$ is also upper-Hessian with the same eigenvalues. There are many ways to achieve such a decomposition but Q must be upper-Hessian too. One way to decompose H is to apply SU(2) rotations in the subspace $\{i, i+1\}$ starting from $i = 1$ and continuing to $i = \dim(H)$.

$$Q = Q_1 \ldots Q_k \tag{B.20}$$

Each matrix Q_i has a form

$$Q_i = \begin{pmatrix} u & v \\ -v^\dagger & u^\dagger \end{pmatrix} \text{ in a subspace } \{i, i+1\} \tag{B.21}$$

with

$$u = h_{i,i}/r \tag{B.22}$$

$$v = -h^\dagger_{i+1,i}/r \tag{B.23}$$

$$r = \sqrt{|h_{i,i}|^2 + |h_{i+1,i}|^2} \tag{B.24}$$

After QR decomposition the system (Eq. B.19) can be transformed

$$MV = VH + R \tag{B.25}$$

$$MVQ = V(QU + \mu)Q + RQ = VH' + R' \tag{B.26}$$

By construction, Q has only two non-zero elements in the last row so that

$$R' = (0, \ldots, r'_{k-1}, r'_k) \tag{B.27}$$

$$r'_{k-1} = r_k Q_{k,k-1} \tag{B.28}$$

$$r'_k = r_k Q_{k,k} \tag{B.29}$$

$$v'_i = \sum_{j=1,i+1} v_j Q_{j,i} \tag{B.30}$$

Note that rotated vectors v'_i satisfy the same system of relations (Eq. B.19) but the residue was multiplied by a coefficient less than one. After rejecting the last vector v_k we are ready to grow the matrix or do QR decomposition again. The order of decomposition was crucial in maintaining the form of the system (Eq. B.19).

Now we have all the parts to the algorithm:

1. Fill the matrix H to level $N > k$, the value $N \geq 2k$ simplifies algorithm.
2. Select $N - k$ shifts μ_i (see below)
3. Do QR decompositions for each μ_i and rotate the vectors. After rotation the subset $V^{(k)}$ and $R^{(k)}$ are ready for the new iteration.
4. If $|r_k| < \epsilon$ all k vectors are lie in the subspace of k eigenvectors with precision ϵ, so the iteration is terminated.

Selection of the shifts μ_i is the crucial and the least defined part of the algorithm. On each iteration of the algorithm above the residual vector r_k is transformed by a polynomial of the operator with roots at μ_i

$$r'_k \leftarrow N_P \prod_i (M - \mu_i) r_k \tag{B.31}$$

with some normalization factor N_P. With a proper choice of shifts one can select the point in eigenvalue space and find k eigenvectors closest to that region (see [76] for more discussion). The simplest choice of shifts comes from the eigenvalues of H itself. The highest $N - k$ eigenvalues of H were chosen as the shifts and the lowest k eigenvalues of H were converging to the lowest eigenvalues of D.

Ideally, the vectors v_i stay orthogonal to each other during successive iterations and the system (Eq. B.19) is always exact. In the real world numerical errors accumulate in algorithm time, especially when using single precision numbers, and extra care should be taken to keep the error under control. The simplest solution is to orthogonalize the vectors v_i at some point in the algorithm. Three levels of orthogonalization were used:

3. Vectors v_i are orthogonalized by the Gramm–Schmidt procedure after rotations.

2. The residue r_N is orthogonalized to v_i after building the matrix $H^{(N)}$.

1. The vectors v_i are re-orthogonalized to all previous vectors as they are computed (Algorithm B.18).

When the operator has almost degenerate eigenvalues higher level of orthogonalization is necessary.

The computational cost of one step of the algorithm consists of k applications of the Dirac operator, $\sim k^2$ vector products and sums, some synchronization time $\sim k$ and the time to compute QR decompositions($\sim k^{3-4}$ but small in our case). A convenient benchmark for the algorithm is one-half cost of the one CG iteration (one application of Dirac operator plus some vector algebra) although different operations can be optimized differently. The results are presented in Figure B-1. The fit curve for region $k \in [4, 48]$ is

$$f(k) = 0.26k^2 + 1.6k + 0.7 \tag{B.32}$$

The total cost of the algorithm behaves differently because the convergence is much better with a bigger vector space. There is no significant difference in convergence time for the number of eigenvectors $k \geq 32$.

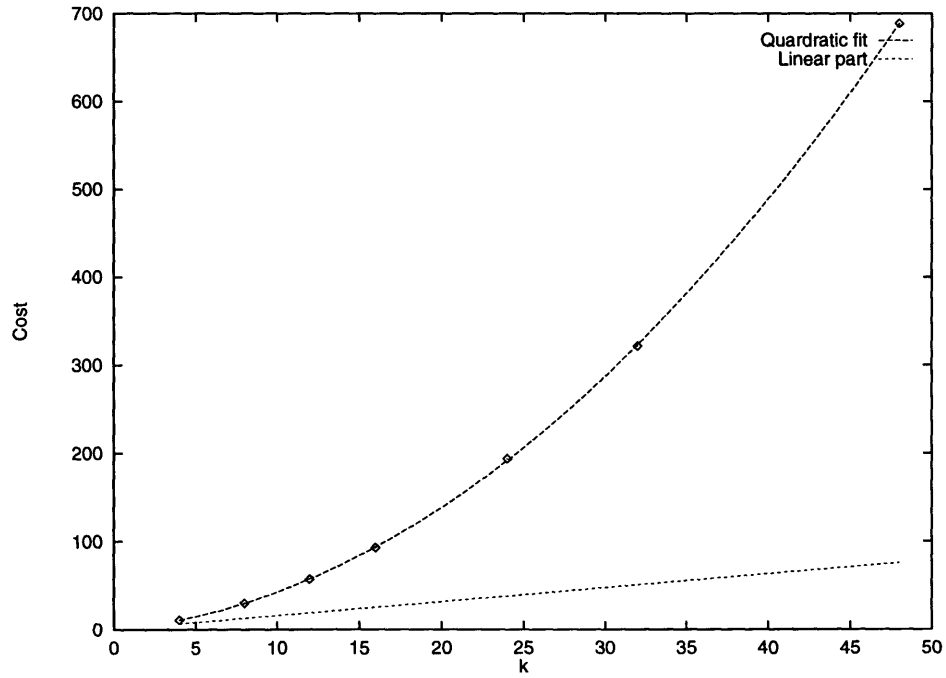

Figure B-1: Dependence of the time of one iterative Arnoldi algorithm iteration on the size of a system

Bibliography

[1] M. F. Atiyah, N. J. Hitchin, V. G. Drinfeld, and Yu. I. Manin. Construction of instantons. *Phys. Lett.*, A65:185, 1978.

[2] M.F. Atiyah and I.M. Singer. The index of elliptic operators. *Ann. Math*, 87:484, 1968.

[3] Ya. Ya. Balitskii and A. V. Yung. Collective – coordinate method for quasizero modes. *Phys. Lett.*, B168:113, 1986.

[4] I. Balitsky and A. Schäfer. Valley method versus instanton-induced effective lagrangian up to $(E/E_{spha})^{8/3}$. *Nucl. Phys.*, B404, 1993.

[5] D. Barkai, K.J.M. Moriarty, and C. Rebbi. Force between static quarks. *Phys. Rev.*, D30, 1984.

[6] G.C. Batrouni, G.R. Katz, A.S. Kronfeld, G.P. Lepage, B. Svetisky, and K.G. Wilson. Langevin simulations of lattice field theories. *Phys. Rev.*, D32, 1985.

[7] A. A. Belavin, A. M. Polyakov, A. S. Shvarts, and Yu. S. Tyupkin. Pseudoparticle solutions of the Yang-Mills equations. *Phys. Lett.*, B59:85, 1975.

[8] C. Bernard, M. Golterman, M. Ogilvie, and J. Potvin, editors. *LATTICE 96: Proc. XIV International Symposium on Lattice Field Theory.* North-Holland, 1997.

[9] K.M. Bitar, R.G. Edwards, U.M. Heller, and A.D. Kennedy. QCD with dynamical Wilson fermions at $\beta = 5.5$. In C. Bernard, M. Golterman, M. Ogilvie, and J. Potvin, editors, *LATTICE 96.* North-Holland, 1997.

[10] M. Bochicchio, L. Maiani, G. Martinelli, G. Rossi, and M. Testa. Chiral symmetry on the lattice with Wilson fermions. *Nucl. Phys.*, B262:331, 1985.

[11] K.C. Bowler, F. Gutbrod, P. Hasenfratz, U. Heller, F. Karsch, R.D. Keway, I. Montvay, G.S. Pawley, J. Smit, and D.J. Wallace. The β-function and potential at $\beta = 6.0$ and $\beta = 6.3$ in SU(3) gauge theory. *Phys. Lett.*, B163:367, 1985.

[12] R. Brower, T. Ivanenko, A. Levi, and K. Orginos. Chronological inversion method for the dirac matrix in Hybrid Monte-Carlo. *hep-lat/9509012*, 1997. To be published.

[13] R.C. Brower, A.R. Levi, and K. Orginos. Extrapolation Methods for the Dirac Inverter in Hybrid Monte Carlo. In *Lattice 94*, volume 42 of *Nucl. Phys. (Proc. Supl)*, 1995.

[14] C.G. Callan, R. Dashen, and D.J. Gross. Toward a theory of the strong interaction. *Phys. Rev.*, D17:2717, 1978.

[15] D.J.E. Callaway and A. Rahman. Lattice gauge theory in microcanonical ensemble. *Phys. Rev.*, D28:1506, 1983.

[16] M.-C. Chu, J.M. Grandy, S. Huang, and J.W. Negele. Correlation functions of hadron currents in the QCD vacuum calculated in lattice QCD. *Phys. Rev.*, D48, 1993.

[17] M.C. Chu, J.M. Grandy, S. Huang, and J.W. Negele. Evidence for the role of instantons in hadron structure from lattice QCD. *Phys. Rev.*, D49, 1994.

[18] M. Creutz. *Quarks, Gluons and Lattices.* Cambridge Univ. Press, 1983.

[19] M. Creutz, editor. *Quantum Fields on the Computer.* World Scientific, 1992.

[20] J. Cullum and R.A. Willoughby. Computing eigenvalues of very large symmetric matrices — an implementation of Lanczos algorithm with no reorthogonalization. *Journal of Comp. Physics.*, 44:329, 1981.

[21] D. Daniel, R. Gupta, G.W. Kilkup, A. Patel, and S.R. Sharpe. Phenomenology with Wilson fermions using smeared sources. *Phys. Rev.*, D46:3130, 1992.

[22] P. de Forcrand, M. Garcia-Perez, and I.-O. Stamatescu. Improved cooling algorithm for gauge theories. In *LATTICE 95.* North-Holland, 1996.

[23] P. de Forcrand, M. Garcia-Perez, and I.-O. Stamatescu. Topology by improved cooling: susceptibility and size distributions. In C. Bernard, M. Golterman, M. Ogilvie, and J. Potvin, editors, *LATTICE 96.* North-Holland, 1997.

[24] P. de Forcrand, M. Garcia-Perez, and I.-O. Stamatescu. Topology of the SU(2) vacuum: a lattice study using improved cooling. *hep-lat/9701012*, 1997. To be published.

[25] P. de Forcrand and S. Kim. Topological susceptibility and instanton size distribution from over-improved cooling. In *LATTICE 95.* North-Holland, 1996.

[26] T. DeGrand and P. Rossi. Conditioning techniques for dynamical fermions. *Comp. Phys. Comm.*, 60, 1990.

[27] S. Duane. Stochastic quantization vs. the microcanonical ensemble: getting the best of both worlds. *Nucl. Phys.*, B257:652, 1985.

[28] S. Duane, A.D. Kennedy, B.J. Pendleton, and D. Roweth. Hybrid Monte Carlo. *Phys. Lett.*, B195:216, 1987.

[29] S. Duane and J. Kogut. Hybrid stochastic differential equations applied to Quantum Chromodynamics. *Phys. Rev. Lett.*, 55, 1985.

[30] S. Duane and J. Kogut. The theory of hybrid stochastic algorithms. *Nucl. Phys.*, B275:398, 1986.

[31] R.G. Edwards, I. Horvath, and A.D. Kennedy. Instabilities and non-reversibility of Molecular Dynamics trajectories. *hep-lat/9606004*, 1996.

[32] R. Frezzotti and K. Jansen. A polynomial Hybrid Monte Carlo algorithm. *hep-lat/9702016*, 1997.

[33] A. Frommer, V. Hannermann, B. Nöckel, T. Lippert, and K. Schilling. Accelerating Wilson fermion matrix inversion by means of the stabilized biconjugate gradient algorithm. *Int. J. Mod. Phys.*, C5:1073, 1994.

[34] A. Frommer, B. Nöckel, S. Güsken, T. Lippert, and K. Schilling. Many masses on one stroke: economic computation of quark propagators. *Int. J. Mod. Phys.*, C6:627, 1995.

[35] M. Garcia-Perez, A. Gonzalez-Arroyo, J. Snippe, and P. van Baal. Instantons from over-improved cooling. *Nucl. Phys.*, B413:535, 1994.

[36] G. Golub and C. Van Loan. *Matrix Computations.* The John Hopkins University Press, Baltimore, 1990.

[37] S. Gottlieb, W. Liu, D. Toussaint, R.L. Renken, and R.L. Sugar. Hybrid Molecular Dynamics algorithms for the numerical simulation of Quantum Chromodynamics. *Phys Rev.*, D35:2531, 1987.

[38] J. Grandy and G. Kilcup. Effect of improving the lattice gauge action on QCD topology. *Nucl. Phys. Proc. Suppl.*, 53:560–563, 1997.

[39] R. Gupta, C.F. Baillie, R.G. Brickner, G.W. Kilkup, A. Patel, and S.R. Sharpe. QCD with dynamical Wilson fermions. II. *Phys. Rev.*, D44, 1991.

[40] R. Gupta, G.W. Kilkup, and S.R. Sharpe. Tuning the Hybrid Monte-Carlo algorithm. *Phys. Rev.*, D38, 1988.

[41] R. Gupta, A. Patel, C.F. Baillie, G. Guralnik, G.W. Kilkup, and S.R. Sharpe. QCD with dynamical Wilson fermions. *Phys. Rev.*, D40, 1989.

[42] M. A. Halasz, T. Kalkreuter, and J. J. M. Verbaarschot. Universal correlations in spectra of the lattice QCD Dirac operator. *Nucl. Phys. Proc. Suppl.*, 53:266–268, 1997.

[43] M. A. Halasz and J. J. M. Verbaarschot. Universal fluctuations in spectra of the lattice Dirac operator. *Phys. Rev. Lett.*, 74:3920–3923, 1995.

[44] C. Hill and J. Marshall. Applications of parallel Navier Stokes solvers to Ocean Modelling. In *Proc. of Parallel CDF 95.* Caltech, 1995.

[45] K. Jansen. Recent developments in fermion simulation algorithms. In C. Bernard, M. Golterman, M. Ogilvie, and J. Potvin, editors, *LATTICE 96.* North-Holland, 1997.

[46] K. Jansen and C. Liu. Study of Liapunov exponents and the reversibility of molecular dynamics algorithms. *Nucl. Phys. (Proc. Suppl.)*, 53:974, 1997.

[47] K. Jansen, C. Liu, H. Simma, and D. Smith. Low-lying eigenvalues of the Wilson-Dirac operator. *Nucl. Phys. Proc. Suppl.*, 53:262–265, 1997.

[48] T. Kalkreuter. Numerical analysis of the spectrum of the Dirac operator in four-dimentional SU(2) gauge fields. *Nucl.Phys.Proc.Suppl.*, 49:168, 1996.

[49] T. Kalkreuter. Study of Cullum's and Willoughby's Lanczos method for Wilson fermions. *Comput.Phys.Comm.*, 95:1, 1996.

[50] T. Kalkreuter and H. Simma. An accelerated conjugate gradient algorithm to compute low-lying eigenvalues – a study for the Dirac operator in SU(2) lattice QCD. *Comp. Phys. Comm.*, 93:33, 1996.

[51] M. Lüscher. A new approach to the problem of dynamical quarks in numerical simulations of lattice QCD. *Nucl. Phys.*, B418:637, 1994.

[52] M. Lüsher. Topology of lattice gauge fields. *Comm. Math. Phys.*, 85, 1982.

[53] R.D. Mawhinney. The status of the Teraflops projects. In *Lattice 94*, volume 42 of *Nucl. Phys. (Proc. Supl.)*, 1995.

[54] G. Parisi and Y. Wu. Perturbation theory without gauge fixing. *Sci. Sin.*, 24:483, 1981.

[55] M. Plagge. Investigation of the biconjugate gradient algorithm for the inversion of fermion matrices. *hep-lat/9212007*, 1992.

[56] J. Polonyi and H.W. Wyld. Microcanonical simulation of fermionic systems. *Phys. Rev. Lett.*, 51:2257, 1983.

[57] A. M. Polyakov. Compact gauge fields and the infrared catastrophe. *Phys. Lett.*, B59:82, 1975.

[58] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical recipes in Fortran: the art of scientific computing.* Cambridge Univ. Press, 1992.

[59] R.C.Brower, T.L. Ivanenko, J.W. Negele, and K.N. Orginos. Instanton distribution in quenched and full QCD. In C. Bernard, M. Golterman, M. Ogilvie, and J. Potvin, editors, *LATTICE 96.* North-Holland, 1997.

[60] T. Schafer, E. V. Shuryak, and J. J. M. Verbaarschot. Baryonic correlators in the random instanton vacuum. *Nucl. Phys.*, B412:143–168, 1994.

[61] R. Setoodeh, C.T.H. Davies, and I.M. Barbour. Wilson fermions on the lattice – a study of the eigenvalue spectrum. *Phys. Lett.*, B213:195, 1998.

[62] E. V. Shuryak and J. J. M. Verbaarschot. QCD instantons at finite temperature. 1. Gluonic interactions and the fermion determinant. *Nucl. Phys.*, B364:255–282, 1991.

[63] E. V. Shuryak and J. J. M. Verbaarschot. Mesonic correlation functions in the random instanton vacuum. *Nucl. Phys.*, B410:55–89, 1993.

[64] E. V. Shuryak and J. J. M. Verbaarschot. Quark propagation in the random instanton vacuum. *Nucl. Phys.*, B410:37–54, 1993.

[65] E. V. Shuryak and J. J. M. Verbaarschot. Random matrix theory and spectral sum rules for the Dirac operator in QCD. *Nucl. Phys.*, A560:306–320, 1993.

[66] E. V. Shuryak and J. J. M. Verbaarschot. Screening of the topological charge in a correlated instanton vacuum. *Phys. Rev.*, D52:295–306, 1995.

[67] E.V. Shuryak. Instantons in QCD(I). *Nucl. Phys.*, B319:521, 1988.

[68] E.V. Shuryak. Instantons in QCD(II). *Nucl. Phys.*, B319:541, 1988.

[69] E.V. Shuryak. Toward the quantitative theory of the instanton liquid. *Nucl. Phys.*, B302:559,574,599,621, 1988.

[70] E.V. Shuryak. Instantons in QCD(III). *Nucl. Phys.*, B328:85, 1989.

[71] E.V. Shuryak. Instantons in QCD(IV). *Nucl. Phys.*, B328:102, 1989.

[72] E.V. Shuryak. Correlation functions in the QCD vacuum. *Rev. Mod. Phys.*, 65:1, 1993.

[73] E.V. Shuryak and T. Schäfer. Instantons in QCD. *hep-ph/9610451*, 1996. To be published.

[74] E.V. Shuryak and J.J.M. Verbaarshot. Chiral symmetry breaking and correlations in the instanton liquid. *Nucl. Phys.*, B341, 1990.

[75] J. Smit and J.C. Vink. Remnants of the index theorem on the lattice. *Nucl. Phys.*, B286:485, 1987.

[76] D.C. Sorensen. Implicit application of polynomial filters in a k-step Arnoldi method. *SIAM J. Matrix Anal. Appl.*, 13:357, 1992.

[77] K. Symanzik. Some topics in quantum field theory. In R. Schrader *et. al.*, editor, *Mathematical problems in theoretical physics*, volume 153 of *Lecture notes in physics*, New York, 1982. Springer.

[78] K. Symanzik. Continuum limit and improved action in lattice theories. *Nucl. Phys.*, B226:187,205, 1983.

[79] G. t'Hooft. Computation of the quantum effects due to a four-dimensional pseudoparticle. *Phys. Rev.*, D14:3432, 1976.

[80] J. J. M. Verbaarschot. On the spectrum of the QCD Dirac operator. *hep-lat/9606009*, 1996.

[81] J. J. M. Verbaarschot. Universal scaling of the valence quark mass dependence of the chiral condensate. *Phys. Lett.*, B368:137–142, 1996.

[82] J.J.M. Verbaarschot. Streamlines and conformal invariance in Yang–Mills theories. *Nucl. Phys.*, B362:33, 1991. ERRATUM-ibid.B386:236,1992.

[83] J.J.M. Verbaarschot. Spectrum of the Dirac operator in a QCD instanton liquid: Two versus three colors. *Nucl. Phys.*, B427:534–544, 1994.

[84] K.G. Wilson. Confinement of quarks. *Phys. Rev.*, D10:2445, 1974.

8050 - 850