Application of the maximum entropy method to dynamical fermion simulations.

Clowser, Jonathan

APPLICATION OF THE MAXIMUM ENTROPY METHOD TO DYNAMICAL FERMION SIMULATIONS.

By

Jonathan Clowser

SUBMITTED IN FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
AT
UNIVERSITY OF WALES SWANSEA
SINGLETON PARK, SWANSEA, SA2 8PP
SEPTEMBER 2002

UNIVERSITY OF WALES SWANSEA

Date: **September 2002**

Author: **Jonathan Clowser**

Title: **Application of the Maximum Entropy Method to Dynamical Fermion simulations.**

Department: **Department of Physics**

Degree: **Ph.D.** Year: **2002**

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signature of Author

29/4/03
Date

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by explicit references. A Bibliography is appended.

Signature of Author

29/4/03
Date

I hereby give consent for my thesis, if accepted, to be available for photocopying and inter-library loan, and for the title and summary to be made available to outside organisations.

Signature of Author

29/4/03
Date

Abstract

This thesis presents results for spectral functions extracted from imaginary-time correlation functions obtained from Monte Carlo simulations using the Maximum Entropy Method (MEM). The advantages this method are (i) no a priori assumptions or parametrisations of the spectral function are needed, (ii) a unique solution exists and (iii) the statistical significance of the resulting image can be quantitatively analysed.

The Gross Neveu model in $d = 3$ spacetime dimensions (GNM_3) is a particularly interesting model to study with the MEM because at $T = 0$ it has a broken phase with a rich spectrum of mesonic bound states and a symmetric phase where there are resonances. Results for the elementary fermion, the Goldstone boson (pion), the sigma, the massive pseudoscalar meson and the symmetric phase resonances are presented.

UKQCD $N_f = 2$ dynamical QCD data is also studied with MEM. Results are compared to those found from the quenched approximation, where the effects of quark loops in the QCD vacuum are neglected, to search for sea-quark effects in the extracted spectral functions. Information has been extract from the difficult axial spatial and scalar as well as the pseudoscalar, vector and axial temporal channels. An estimate for the non-singlet scalar mass in the chiral limit is given which is in agreement with the experimental value of $M_{a_0} = 985$ MeV.

Declaration

This thesis has been written and composed by me and contains the results of my work except where explicitly stated by references. The derivations which appear in § 4.4 were the work of my collaborator Dr. S. J. Hands.

The results in Chapter 4 appear in:

- J. E. Clowser and C. G. Strouthos, Nucl. Phys. B (Proc. Suppl.) **106** (2002), 489-491
- C. R. Allton, J. E. Clowser, S. J. Hands, J. B. Kogut and C. G. Strouthos, hep-lat/0208027, submitted to Phys. Rev. D

and the results in Chapter 5 are a result of my work as part of the UKQCD collaboration and will appear in:

- C. R. Allton and J. E. Clowser, in preparation

Other publications to which I have contributed

- C. R. Allton, D. Blythe and J. E. Clowser, Nucl. Phys. B (Proc. Suppl.) **109** (2002), 192-196
- C. R. Allton, D. Blythe and J. E. Clowser, in preparation

Signature of Author

Acknowledgements

I would like to thank my family, and my girlfriend Karen for the past few years for all their love and support throughout my time at University. Thanks to all my friends who have made my time in Swansea one I shall look back on with the fondest of memories.

Many thanks to my supervisor Chris Allton for all his help over the course of my Ph.D. and for the thorough proof reading of this thesis. I would also like to thank Simon Hands and Costas Strouthos for many helpful discussions on the work included in Chapter 4. Thanks also to UKQCD Collaboration for the use of their dynamical QCD data used for Chapter 5.

Contents

1 Introduction **1**

1.1 Lattice QCD 2

1.1.1 Path integral formalism of QCD 2

1.2 The continuum QCD action 4

1.3 Lattice gauge theory formalism 6

1.3.1 Spacetime discretisation 6

1.3.2 Lattice representation of the fermion and gauge fields 7

1.3.3 Discretisation of the action 9

1.3.4 The integration measure 10

1.3.5 Numerical simulation 11

1.4 Simulating QCD 13

1.4.1 The quenched approximation 13

1.4.2 $N_f = 2$ dynamical quarks 15

1.5 $\mathcal{O}(a)$ improvement 16

1.6 The Gauge action 16

1.7 The fermion action 17

1.8 Hadronic correlation functions 18

1.9 Why MEM? 19

1.10 Overview of thesis 21

2 Maximum Entropy Method **23**

2.1 Bayes' Theorem 23

2.2 The likelihood function 24
2.2.1 Normalisation 26
2.2.2 Comparison to standard χ^2-fitting 26
2.3 The prior probability 26
2.3.1 Discretisation of frequency, ω 27
2.3.2 Normalisation 27
2.4 MEM algorithm 28
2.4.1 Step 1 - Find the most probable $f(\omega)$ for a given α 28
2.4.2 Step 2 - Calculate final image 29
2.4.3 Step 3 - Error analysis 30
2.4.4 Condition for integrating over α 31
2.5 Maximising Q 32
2.5.1 Extremum condition 32
2.5.2 SVD of lattice kernel 33
2.5.3 Bryan's algorithm 34
2.5.4 Newton search for solution 35
2.5.5 Diagonalising Eq. 2.50 36
2.5.6 Getting $\delta\mathbf{u}$ from $Y^{-1}\delta\mathbf{u}$ 38
2.5.7 How to deal with small/negative ξ's 38
2.5.8 Diagonalising Eq. 2.62 40
2.6 Uniqueness of the solution in MEM 42

3 Testing the Maximum Entropy Method **45**
3.1 Test data generation 46
3.1.1 Delta function 46
3.1.2 QCD-like data 47
3.2 The tests 49
3.2.1 Simple pole analysis 49
3.2.2 Input data quality 50
3.2.3 Changing the default model 51

3.2.4 Changing the prior probability for α 52
3.2.5 Spectral width 52
3.3 Summary 53

4 Modelling QCD: The Four-Fermion Interaction 55
4.1 Essential properties 57
4.2 Mean field analysis at zero temperature and density 60
4.3 Simulations 63
4.3.1 Data sets 66
4.4 Theoretical spectral function form 66
4.4.1 Broken phase π, f and PS channels 66
4.4.2 Symmetric phase 68
4.4.3 Broken phase sigma 69
4.5 Results 71
4.5.1 Broken phase π, f and PS channels 71
4.5.2 Symmetric phase 75
4.5.3 Broken phase sigma 82
4.6 Conclusions 86

5 Dynamical QCD 88
5.1 Simulation parameters 88
5.2 MEM technicalities 91
5.3 Results 91
5.3.1 Quenched data set 92
5.3.2 Matched Unitary Set 95
5.3.3 Lightest κ_{sea} data 100
5.4 Non-singlet scalar meson 101
5.5 Conclusions 105

6 Conclusions 107
6.1 Summary of the GNM_3 analysis (Chapter 4) 107

6.2 Summary of QCD analysis (Chapter 5) 108
6.3 Future work 109
6.3.1 GNM_3 109
6.3.2 Dynamical QCD 109

A Entropy definition **110**
A.1 Monkey argument for entropy and prior probability 110
A.2 Axiomatic construction of entropy 112

B Singular value decomposition **116**

Bibliography **119**

Chapter 1

Introduction

Quantum Chromodynamics (QCD) is well established as the theory of the strong interaction. This has been due to the success of perturbative methods in describing the short distance (large momentum transfer) behaviour of quarks and gluons. However, QCD also needs to exhibit experimentally observed phenomenon, such as quark confinement (i.e. quarks bound in colour singlet states). This necessitates calculating at low energies, $\mu \lesssim 1\text{GeV}$, associated with the characteristic length scale of a hadron of approximately 1fm. Since QCD exhibits the property of asymptotic freedom, at such energies, the strong coupling becomes of $\mathcal{O}(1)$. Thus perturbative methods fail, and a non-perturbative approach, such as Lattice QCD, is required.

Lattice QCD was originally formulated in 1974 by K. G. Wilson [1]. It provides a non-perturbative mechanism for confinement in the strong coupling limit and enables numerical studies of the low energy behaviour of QCD. The first numerical results were reported in [2, 3]. Reproducing the experimentally observed hadron spectrum validates both QCD and the lattice approach, and confidence in the results obtained means that it can be used to predict other phenomenology which cannot be measured directly from experiment. Due to the considerable computer time required to simulate

QCD, the quenched approximation (in which the quark loops in the vacuum are neglected) has been used, and in fact is still used to this day. The applicability of this approximation can be measured by comparing the light hadron spectrum with experiment. In order to make precise comparisons it is necessary to have good control over systematic effects such as lattice artifacts. Comparisons of this type all seem to agree that the quenched light hadron spectrum agrees with experiment at the 10% level [4, 5].

Now with the recent theoretical developments and additional computational power, it is possible to simulate full QCD (i.e. include quark loops in the vacuum). Although the simulations have not reached the stage where precision measurements of the spectrum can be made it is hoped that the 10% discrepancy seen in the quenched approximation is reduced. It is also interesting to see if the effects of these dynamical quarks can be seen, for example the decay of $\rho \rightarrow \pi\pi$. Using the Maximum Entropy Method (MEM) the spectral function of the particles can be extracted directly from the lattice correlation data. This method is described in detail in Chapter 2.

1.1 Lattice QCD

In this section a brief outline of the fundamental elements of lattice QCD is given. For more detailed information see the textbooks [6, 7] or introductory lecture courses [8, 9].

1.1.1 Path integral formalism of QCD

All information about the physical observables in the theory are contained within an infinite number of vacuum expectation values of time-ordered products of quantum

field operators, known as Green's functions. Using the path integral formalism [10], these quantum probability amplitudes can be directly related to the probability distribution of the classical fields

$$\langle 0|\hat{\mathcal{T}}\{\hat{\mathcal{O}}[\bar{\psi},\psi,A_\mu]\}|0\rangle = \frac{1}{Z}\int \mathcal{D}\bar{\psi}\mathcal{D}\psi\mathcal{D}A_\mu \mathcal{O}[\bar{\psi},\psi,A_\mu]e^{iS[\bar{\psi},\psi,A_\mu]}, \tag{1.1}$$

where $\hat{\mathcal{O}}[\bar{\psi},\psi,A_\mu]$ corresponds to a product of quantum operators , $\mathcal{O}[\bar{\psi},\psi,A_\mu]$ to a product of fermion, anti-fermion and gauge fields. The time ordering of the operators is denoted by $\hat{\mathcal{T}}$ and S is the classical action. The partition function, Z, is defined as

$$Z = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\mathcal{D}A_\mu e^{iS[\bar{\psi},\psi,A_\mu]}. \tag{1.2}$$

Note that here and throughout this thesis the usual nomenclature, known as natural units, $c = \hbar = 1$ has been used.

The functional integral in Eq. 1.1 is extremely hard to evaluate numerically due to it being both complex and strongly oscillating. To deal with these problems and make the numerical simulation much easier a Wick rotation, $x_0 \to -ix_4$ can be used to analytically continue from Minkowski to Euclidean spacetime. The partition function is altered to

$$Z = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\mathcal{D}A_\mu e^{-S_E[\bar{\psi},\psi,A_\mu]}. \tag{1.3}$$

where S_E is the Euclidean action (see § 1.2 for definition). The weight, e^{-S_E} in the

partition function is now very similar to that used in Statistical Mechanics, provided that S_E is both real and bounded from below.

1.2 The continuum QCD action

The QCD action is invariant under local $SU(3)$ gauge transformations, $G(x)$. The fermion and gauge fields transform as

$$\psi(x) \rightarrow G(x)\psi(x), \tag{1.4}$$

$$\bar{\psi}(x) \rightarrow \bar{\psi}(x)G^{-1}(x), \tag{1.5}$$

$$A_\mu(x) \rightarrow G(x)A_\mu(x)G^{-1}(x) - (\partial_\mu G(x))G^{-1}(x). \tag{1.6}$$

Thus the continuum QCD action in four-dimensional Euclidean spacetime is given by

$$S_E(x) = \int d^4x \frac{-1}{2g_0^2}\mathrm{Tr}(F_{\mu\nu}(x)F_{\mu\nu}(x)) + \sum_{f=1}^{N_f} \bar{\psi}_f(x)(\not{D} + m_f)\psi_f(x), \tag{1.7}$$

where N_f is the number of fermion flavours and $\mu, \nu = 1 \ldots 4$. The first term describes the dynamics of the gluon gauge fields, A_μ. The gauge field strength tensor, $F_{\mu\nu}(x)$, is defined in terms of the commutator of the covariant derivative, $D_\mu = \partial_\mu + A_\mu(x)$, as

$$F_{\mu\nu}(x) = [D_\mu, D_\nu] = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) + [A_\mu(x), A_\nu(x)] \tag{1.8}$$

The eight gluon gauge fields are defined in terms of the generators of $SU(3)$, T.

$$A_\mu^\dagger(x) = -A_\mu(x), \quad A_\mu(x) = -ig_0 A_\mu^a(x) T^a, \quad a = 1, \ldots, 8. \tag{1.9}$$

The following commutation relations and normalisation conditions are satisfied by the generators

$$[T^a, T^b] = i f_{abc} T^c, \quad \mathrm{Tr}[T^a, T^b] = \frac{1}{2}\delta_{ab}. \tag{1.10}$$

where f_{abc} are the anti-symmetric structure constants and the generators are represented by the eight Gell-Mann matrices, $T^a = \lambda^a/2$.

The second term in Eq. 1.7 is the Euclidean Dirac action describing the interaction of the fermion fields. The Dirac spinor and colour indices have been suppressed and m_f is the mass of the fermion with flavour f. Note that $\not{D} \equiv \gamma_\mu D_\mu$. From now on the sum over the flavours, written in explicitly in Eq. 1.7, will be omitted. The Euclidean Dirac matrices, γ, are related to the Minkowski matrices, γ^M, by

$$\gamma_4 = \gamma_0^M, \quad \gamma_j = -i\gamma_j^M, \quad j = 1, 2, 3, \tag{1.11}$$

and satisfy the condition of Hermiticity and commutation relations below.

$$\gamma_\mu^\dagger = \gamma_\mu, \quad \{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}. \tag{1.12}$$

This action can now be inserted into the path integral formalism. Due to the gauge invariance of the action, this path integral is not well defined, hence the integration is performed over an infinite number of physically equivalent gauge configurations. This problem can be solved by discretising spacetime on a four dimensional lattice.

1.3 Lattice gauge theory formalism

The original formulation of lattice gauge theory was proposed in [1]. A detailed description of the steps involved in discretising the continuum theory, outlined briefly below, can be found in [6, 7].

1.3.1 Spacetime discretisation

Spacetime is discretised onto a four dimensional isotropic hypercubic lattice, Λ_E

$$\Lambda_E = \{x \in \mathbb{R}^4 | x_\mu/a \in \mathbf{Z},\ \mu = 1, \ldots, 4\}, \tag{1.13}$$

where a is the lattice spacing. The integration in the action is now replaced by the sum over all lattice sites, x,

$$\int d^4x \to a^4 \sum_x . \tag{1.14}$$

All dimensionful variables are re-scaled by the lattice spacing, a, to yield dimensionless quantities, e.g. the fermion mass m is replaced by am, since in natural units ($c = \hbar = 1$) mass has the dimensions of inverse length.

1.3.2 Lattice representation of the fermion and gauge fields

The fermion fields are represented by anti-commuting Grassmann variables on the lattice sites. Since the action is bi-linear in the quark fields, the integration over the fermion variables can be performed analytically.

The representation of the gauge fields on the lattice is not so straight forward. If field variables situated on the lattice sites are used then the gauge invariance is spoiled due to the discretisation of the derivative by a finite difference. To retain the gauge invariance of the theory the following procedure can be used.

In the presence of a gauge field in the continuum, a quark field transported from x to y gains a phase factor

$$\psi(y) = \mathcal{P}exp\left\{-\int_x^y A_\mu(z)dz_\mu\right\}\psi(x) \equiv U(y,x)\psi(x) \tag{1.15}$$

where $\mathcal{P}$ denotes a path ordered product, which is required due to the non-abelian nature of the gauge fields. Under a local $SU(3)$ gauge transformation, the parallel transporter, $U(y,x)$, transforms as

$$U(y,x) \rightarrow G(y)U(y,x)G^{-1}(x), \tag{1.16}$$

hence

$$\bar{\psi}(y)U(y,x)\psi(x) \tag{1.17}$$

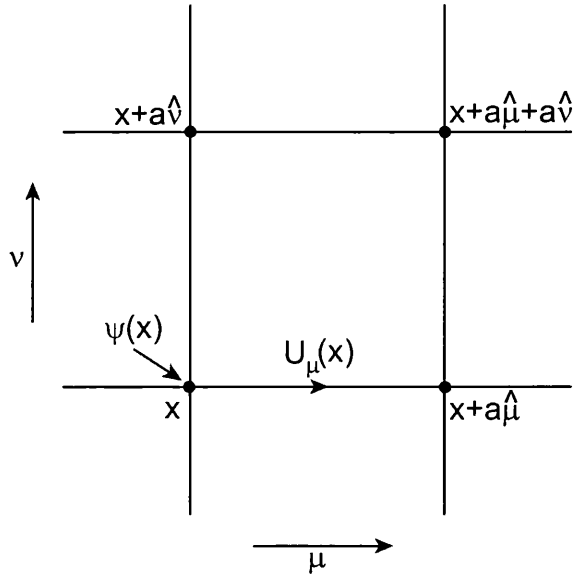

Figure 1.1: Lattice variables, quarks are defined on the lattice sites, gluons on the links.

is gauge invariant. This parallel transporter is represented by a link variable, $U_\mu(x)$, from lattice site x, oriented in the direction μ.

$$U_\mu(x) = e^{-aA_\mu(x+\frac{a\hat{\mu}}{2})}, \tag{1.18}$$

where $\hat{\mu}$ is a unit vector in the direction μ. From the path ordering condition it follows that $U_\mu(x) = U_{-\mu}(x + a\hat{\mu})$. These link variables are represented by 3×3 unitary matrices with unit determinant in the fundamental representation of $SU(3)$. They transform as

$$U_\mu(x) \to G(x)U_\mu(x)G^{-1}(x + a\hat{\mu}), \tag{1.19}$$

provided $G(x)$ belongs to the same representation of the group as the gauge links.

Two types of gauge invariant objects can now be formed on the lattice, this can be seen from Eqs. 1.4, 1.5 and 1.19. Firstly a string defined by

$$\bar{\psi}(y)U_{\mu}(y)\ldots U_{\nu}(x-a\hat{\nu})\psi(x), \tag{1.20}$$

where the trace over the colour indices is implicit and the gauge links are path ordered. Secondly the trace of a product of gauge links forming a closed loop, referred to as a Wilson loop. The simplest example of this is the plaquette, $\mathrm{Tr}U_{\Box}$, where

$$U_{\Box} \equiv U_{\mu}(x)U_{\nu}(x+a\hat{\mu})U_{\mu}^{\dagger}(x+a\hat{\nu})U_{\mu}^{\dagger}(x) \tag{1.21}$$

is the product of links around an elementary square on the lattice.

1.3.3 Discretisation of the action

The QCD action could be discretised in many ways, but a vital condition to be satisfied is that the continuum action must be reproduced in the limit where the lattice spacing goes to zero. This means higher order terms can be added into the lattice action so long as they disappear in the continuum limit.

The lattice action is constructed from the gauge invariant quantities defined in § 1.3.2.

$$S[\bar{\psi},\psi,U] = S_G[U] + S_F[\bar{\psi},\psi,U], \tag{1.22}$$

where S_G is a pure gauge term and S_F depends on the fermionic fields and is defined as

$$S_F[\bar{\psi}, \psi, U] = \sum_{x,y} \bar{\psi}(x) M_{x,y}[U] \psi(y), \tag{1.23}$$

where M is the fermion matrix.

1.3.4 The integration measure

The partition function can now be expressed in terms of the lattice variables

$$Z = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\mathcal{D}U e^{-S_G[U] - \sum_{x,y} \bar{\psi}(x) M_{x,y}[U] \psi(y)}, \tag{1.24}$$

where

$$\mathcal{D}\bar{\psi}\mathcal{D}\psi \equiv \prod_x d\bar{\psi}(x) d\psi(x), \quad \mathcal{D}U \equiv \prod_{x,\mu} dU_\mu(x). \tag{1.25}$$

After performing the integration over the Grassmann valued fermion fields analytically the partition function becomes

$$Z = \int \prod_{x,\mu} dU_\mu(x) \det M[U] e^{-S_G[U]}. \tag{1.26}$$

dU, defined by the conditions below, is known as the Haar measure [6] and is gauge invariant.

$$\int_G dU f(U) = \int_G dU f(UV) = \int_G dU f(VU), \tag{1.27}$$

where $V \in SU(3)$ and $f(U)$ is an arbitrary function over the group. Since the gauge links are elements of a compact group the normalisation condition

$$\int_G dU = 1 \tag{1.28}$$

can be imposed, which reduces the path integral down to a large, but finite, number of integrations and removes the need for gauge fixing. The remaining integration, over the gauge links, is performed numerically.

1.3.5 Numerical simulation

In terms of lattice variables, the physical observables, expressed in the path integral formalism become

$$\langle 0|\hat{\mathcal{T}}\{\hat{\mathcal{O}}(\bar{\psi},\psi,A_\mu)\}|0\rangle = \frac{1}{Z}\int \mathcal{D}\bar{\psi}\mathcal{D}\psi\mathcal{D}U\mathcal{O}(\bar{\psi},\psi,U)e^{-S[\bar{\psi},\psi,U]}. \tag{1.29}$$

There are only a few types of terms which do not vanish, those which contain equal numbers of fermion and anti-fermion fields, due to the Grassmann integration rules [6]. So for a general operator

$$\int \mathcal{D}\bar{\psi}\mathcal{D}\psi e^{-\sum_{x,y}\bar{\psi}(x)M_{x,y}[U]\psi(y)} = \det M[U] \tag{1.30}$$

and

$$\int \mathcal{D}\bar{\psi}\mathcal{D}\psi\psi(x')\bar{\psi}(y')e^{-\sum_{x,y}\bar{\psi}(x)M_{x,y}[U]\psi(y)} = M^{-1}_{x',y'}\det M[U]. \tag{1.31}$$

Dividing Eq. 1.31 by Eq. 1.30 gives the quark propagator in a background gauge field, $G^{\alpha,\beta}_{a,b}(x,y;U)$ in terms of the inverse fermion matrix

$$G^{\alpha,\beta}_{a,b}(x,y;U) = M^{-1}_{\alpha,a,x;\beta,b,y}[U], \tag{1.32}$$

where (a,b) are the colour indices and (α,β) are the Dirac spinor indices. It is from this propagator that the correlation functions are constructed.

Once the fermionic part of the integration has been performed analytically, the expectation value of an operator is given by a path integral over the gauge fields

$$\langle 0|\hat{\mathcal{T}}\{\hat{\mathcal{O}}(\bar{\psi},\psi,U)\}|0\rangle = \frac{1}{Z}\int \mathcal{D}U\mathcal{O}(U,M^{-1}[U])e^{-S_{\text{eff}}[U]}, \tag{1.33}$$

where S_{eff} is the effective action

$$S_{\text{eff}} = S_G[U] - \ln\det M[U]. \tag{1.34}$$

Provided S_{eff} is real valued, the remaining integration over the gauge links can be performed numerically. This is done using Monte Carlo techniques, by generating gauge field configurations $\{U\}_i$ (assignment of a link variable to every link on the

lattice) with a probability proportional to $e^{-S_{\text{eff}}[U]}$. Subsequent configurations are then generated via an algorithm and are separated by several algorithmic updates, known as sweeps or trajectories. This is to reduce the correlations between subsequent configurations. An observable is then calculated as the ensemble average of the observable measured on each configuration.

$$\langle 0|\hat{\mathcal{T}}\{\hat{\mathcal{O}}[\bar{\psi},\psi,A_\mu]\}|0\rangle \approx \frac{1}{N}\sum_{i=1}^{N}\mathcal{O}(\{U\}_i, M^{-1}[\{U\}_i]) \tag{1.35}$$

The statistical error in such an average is $1/\sqrt{N}$ for independent configurations, for correlated configurations this estimate is increased.

1.4 Simulating QCD

To achieve reliable results with acceptable statistical errors from QCD simulations, a sizeable computational effort is required. In order to achieve this an approximation, known as the quenched approximation, is often used. One of the main aims of MEM is to be able to see unquenching effects, i.e. differences between quenched data and dynamical data.

1.4.1 The quenched approximation

The most intensive part in a QCD simulation is calculating the determinant of the fermion matrix, M, in the effective action (Eq. 1.34). This is due to the considerable size of the fermion matrix ($(4_{\text{spins}} \times 3_{\text{colours}} \times N_{\text{lattice sites}})^2$ elements) and the non-local nature of the inverse of M, which is required for the algorithmic update of the configurations. If the approximation

$$\det M[U] = \text{const} \quad \text{(usually set to 1)} \tag{1.36}$$

is made then the computational overhead can be significantly reduced. This is known as the *quenched approximation* and corresponds to neglecting the quark–anti-quark loops in the vacuum (i.e. these quarks, commonly called *dynamical fermions*, are made infinitely heavy and therefore decouple from the theory).

The quenched theory of QCD retains most of the important features seen in full QCD such as confinement and chiral symmetry breaking, but there are consequences. For example, a resonance such as the ρ meson become stable, in full QCD the ρ receives a contribution from an intermediate state, consisting of 2 π's. This is one of the signals MEM will seek, a peak corresponding to two pions in the spectral function.

One of the main effects of quenching is to shift the coupling, this means that quenched and dynamical simulations should not be compared at the same value of the coupling, rather at the same lattice spacing. Such data sets are known as a "matched" ensemble [11].

Although the quenched approximation has yielded good results for observables such as the light hadron spectrum (i.e. to an accuracy of 10% with experimental results), with the advancement in computer power, along with algorithmic development, it is now feasible to simulate two light flavours ($N_f = 2$) of dynamical quarks.

1.4.2 $N_f = 2$ dynamical quarks

The effective action in Eq. 1.34 needs to be real so that importance sampling can be easily implemented. Hence $\det M[U]$ needs to be real and positive. To check that the determinant of the fermion matrix is real is relatively simple, it follows from the lattice Hermiticity relation

$$M[U] = \gamma_5 M^\dagger[U] \gamma_5. \tag{1.37}$$

This does not guarantee the positivity of the determinant though. For example consider an action with two fermion terms (u and d). When the integration over the fermion fields is performed the following determinant is obtained

$$\det M[U] \rightarrow \det M_u[U] \det M_d[U] = (\det M_u[U])^2 \geq 0 \tag{1.38}$$

Where the last statement of positivity only holds if $M_u = M_d$. Hence in the case of pairs of degenerate quarks, positivity is guaranteed.

Whether simulating using the quenched approximation or full QCD, the inversion of the fermion matrix is necessary. This can be very computationally intensive, especially when using the physical masses of the light quarks. It is for this reason that the quark propagators are generated at heavy quark masses, and then the lattice masses calculated are extrapolated to the physical light quark masses.

1.5 $\mathcal{O}(a)$ improvement

The discrete nature of the lattice is one of the major sources of error in lattice simulations. Due to the computational cost of simulating at an arbitrarily small value of the lattice spacing, simulations must be run at small, yet finite lattice spacings. For $N_f = 2$ dynamical fermion simulations, which already have high computational costs the lattice used must be even coarser than for the quenched approximation, $a \approx 0.1$fm. As a consequence of this, the discretisation errors become larger, hence the need to use an $\mathcal{O}(a)$ improved action.

1.6 The Gauge action

Although the gauge action can be improved, the data analysed in Chapter 5 of this thesis is generated using the standard Wilson gauge action [1] defined in Eq. 1.39

$$S_G[U] = \beta \sum_{\square} \left(1 - \frac{1}{3}\Re\mathrm{Tr}U_{\square}\right), \quad \beta = \frac{6}{g_0^2} \tag{1.39}$$

where β parameterises the dependence on the strong coupling constant. The sum is over all positively orientated plaquettes $U_{\square}$ defined in Eq. 1.21. Substituting Eq. 1.18 into $S_G[U]$, the Yang-Mills term in the continuum action is obtained up to discretisation errors of $\mathcal{O}(a^2)$. Improved gauge actions can be used, however. The CP-PACS collaboration use an improved gauge action given in Eq. 1.40 for their "full" QCD simulations [12]

$$S_G^I[U] = \beta \left(\sum_{1\times1} W_{1\times1} - 0.0907 \sum_{1\times2} W_{1\times2}\right), \quad W_{1\times1} = \Re\mathrm{Tr}U_{\square} \tag{1.40}$$

where $W_{1\times 2}$ denotes the real part of a trace over a 1×2 rectangular Wilson loop. The sums are always over positively orientated loops. The motivation and method used in obtaining this action can be found in [13] and will not be discussed further in this thesis.

1.7 The fermion action

The fermion action used for the QCD work in Chapter 5 is the $\mathcal{O}(a)$ improved Wilson fermion action

$$S_F[\bar{\psi}, \psi, U] = S_F^{\text{Wilson}}[\bar{\psi}, \psi, U] + S_F^{SW}[\bar{\psi}, \psi, U]. \tag{1.41}$$

The first term is the standard Wilson fermion action [14] and the second term is a counterterm, known as the Scheikoleslami-Wohlert or Clover term [15] which can be tuned in order to cancel to $\mathcal{O}(a)$ discretisation errors arising from the Wilson fermion action. The discretised form is

$$S_F^{SW}[\bar{\psi}, \psi, U] = -c_{sw}\frac{i\kappa}{2}\sum_{x,y,\mu}\bar{\psi}(x)\sigma_{\mu\nu}F_{\mu\nu}(x)\psi(x) \tag{1.42}$$

where $\sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu]$ and $F_{\mu\nu}$, the lattice field strength tensor, is defined by

$$F_{\mu\nu}(x) = \frac{1}{8}(Q_{\mu\nu}(x) - Q_{\mu\nu}^\dagger(x)) \tag{1.43}$$

where $Q_{\mu\nu}$ is the sum of the four plaquettes around lattice site x in the μ, ν plane

$$\begin{aligned} Q_{\mu\nu} = \quad & U_\mu(x)U_\nu(x+\hat{\mu})U_\mu^\dagger(x+\hat{\nu})U_\nu^\dagger(x) \\ + \quad & U_\nu(x)U_\mu^\dagger(x-\hat{\mu}+\hat{\nu})U_\nu^\dagger(x-\hat{\mu})U_\mu(x-\hat{\mu}) \\ + \quad & U_\mu^\dagger(x-\hat{\mu})U_\nu^\dagger(x-\hat{\mu}-\hat{\nu})U_\mu(x-\hat{\mu}-\hat{\nu})U_\nu(x-\hat{\nu}) \\ + \quad & U_\nu^\dagger(x-\hat{\nu})U_\mu(x-\hat{\nu})U_\nu(x+\hat{\mu}-\hat{\nu})U_\mu^\dagger(x). \end{aligned} \tag{1.44}$$

The $\mathcal{O}(a)$ improved action is then

$$S_F[\bar{\psi},\psi,U] = \sum_{x,y} \bar{\psi}(x) M_{x,y}[U] \psi(y), \tag{1.45}$$

with the fermion matrix, $M_{x,y}[U]$, defined by

$$\begin{aligned} M_{x,y}[U] = \quad & \delta_{x,y}\left(1 - c_{sw}\frac{i\kappa}{2}\sum_{\mu\nu}\sigma_{\mu\nu}F_{\mu\nu}(x)\right) \\ - \quad & \kappa\sum_{\mu}[\delta_{x+\hat{\mu},y}(1-\gamma_\mu)U_\mu(x) + \delta_{x-\hat{\mu},y}(1+\gamma_\mu)U_\mu^\dagger(y)] \end{aligned} \tag{1.46}$$

c_{sw} is known as the clover coefficient and is a function of the bare coupling g_0. It can be tuned in order to remove $\mathcal{O}(a)$ discretisation errors.

1.8 Hadronic correlation functions

The interpolating operator, $J(x)$, for a meson is given by

$$J(x) = \bar{\psi}(x)\gamma_\mu\gamma_5\psi(x). \tag{1.47}$$

The (zero-momentum) hadronic correlation function for mesons is defined as

$$G_2(t) = \sum_{\{U\}} \sum_{\mathbf{x}} \langle 0|J(\mathbf{x},t)J^\dagger(\mathbf{0},0)|0\rangle \tag{1.48}$$

Introducing a full set of states this becomes

$$G_2(t) = \sum_{\{U\},\mathbf{x}} \sum_i \int \frac{d^3k}{2E} \langle 0|J(\mathbf{x},t)|P_i(\mathbf{k})\rangle\langle P_i(\mathbf{k})|J^\dagger(\mathbf{0},0)|0\rangle \tag{1.49}$$

and performing the sum over $\mathbf{x}$

$$G_2(t) = \sum_{\{U\}} \sum_i \frac{1}{2M_i} \langle 0|J(0)|P_i(0)\rangle\langle P_i(0)|J^\dagger(0)|0\rangle e^{-M_i t}. \tag{1.50}$$

Eq. 1.50 is a sum of exponentials, one for each state in the channel each with a different mass M_i. The largest mass will decay fastest and the lightest (ground) state will decay the slowest. In the standard χ^2-fitting method, only large times, when the excited states have all decayed away and only the ground state remains are fitted to.

$$\lim_{\text{t large}} G_2(t) = \frac{|\langle 0|J(0)|P(0)\rangle|^2}{2M_0} e^{-M_0 t}. \tag{1.51}$$

1.9 Why MEM?

The hadronic spectral functions in QCD play a vital role in understanding all the properties of hadrons and the QCD vacuum structure. For example, consider e^+e^-

annihilation into hadrons, this can be expressed in terms of the spectral function corresponding to the correlation function of the QCD electromagnetic current.

Numerical simulations of QCD have so far been very successful at extracting "static" properties such as hadron masses and decay constants. It is desirable to extend its powers to the extraction of the hadronic spectral functions. However, there is much difficulty in accessing the "dynamical" quantities, such as spectral functions, from the finite set of discrete points in imaginary time generated by the Monte Carlo simulations. The analytic continuation from imaginary to real time using the limited and noisy lattice data available is a typical ill-posed problem, where the number of data points is much smaller than the degrees of freedom to be reconstructed. The standard method (χ^2-fitting) is clearly inapplicable here, since many degenerate solutions would be found. This is why the first attempts at extracting spectral functions relied on fitting to a specific ansätze [16, 17, 18, 19, 20, 21]. There are two major drawbacks to these previous approaches

- a priori assumptions for the spectral shape prevent the study of the fine structures contained in it, and
- the result does not remain stable under a change in the number of parameters used in the specific spectral function ansätze used.

Both of these problems become even more severe at finite temperature where very little is known about the spectral shape.

The maximum entropy method provides a way to extract the spectral functions from the lattice correlation data in which Shannon's information entropy [22] plays

a vital role. The first application of information entropy to statistical mechanics was made by Jaynes [23] and to optical image reconstruction by Frieden [24]. Since then MEM has been applied to many different scientific fields including analysis of quantum Monte Carlo simulations in condensed matter physics [25, 26, 27, 28, 29, 30, 31] and image reconstruction for crystallography [32] and astrophysics [33]. In the context of QCD, MEM uses Bayesian probability theory [34] to make a statistical inference on the most probable spectral function from a given Monte Carlo data set. A priori assumptions about the functional form are not made, nevertheless, for any given lattice data, a unique solution exists. Furthermore, error analysis can be carried out on the resultant image so the statistical relevance of any feature can be evaluated. This method opens up a whole host of possibilities for further study beyond the conventional methods of fitting lattice data.

1.10 Overview of thesis

In chapter 2 the MEM algorithm used throughout this thesis to obtain the images will be discussed in detail. Chapter 3 presents some results from runs on test data, showing that the method works for a variety of cases. Also the image quality and variation will be tested by varying a number of factors including the data quality (i. e. number of time slices, noise levels and number of configurations). In chapter 4 the results from a model of QCD are presented, preceded by some basic theory for the model and the expected forms for the spectral functions. This is the first attempts at applying MEM to data with dynamical fermions. Chapter 5 presents results obtained from UKQCD $N_f = 2$ dynamical data as well as a quenched data set for reference. This is the first application of MEM beyond the quenched approximation. Finally in

Chapter 6 a summary of the conclusions from this results presented in this thesis is given.

Chapter 2

Maximum Entropy Method

In the chapter a detailed discussion on the maximum entropy method (MEM) will be given. Beginning with the foundation of Bayes' theorem and going onto a detailed proof of Bryan's algorithm [35], the method used for all the results presented in this thesis.

2.1 Bayes' Theorem

The theoretical basis for MEM is Bayes' theorem in probability theory [34]. This states that

$$P[X|Y] = \frac{P[Y|X]P[X]}{P[Y]}, \tag{2.1}$$

where $P[X|Y]$ is known as the conditional probability of X given Y, i.e. the probability that event X occurs given that event Y has already happened. $P[X]$ is just the probability that event X occurs independent of event Y.

Now re-write this in terms of the Monte Carlo lattice data, D, and the spectral function, f. Included also in the following expression is a hypothesis term, H, which

represents all the a priori knowledge (e.g. $f(\omega \geq 0) \geq 0$), i.e. the spectral function for a particle is positive semi-definite and only defined for positive energies). Bayes' theorem now reads

$$P[f|DH] = \frac{P[D|fH][P[f|H]}{P[D|H]}. \tag{2.2}$$

The terms on the right hand side of Eq. 2.2 are known as:

$P[D\|fH]$	-	likelihood function
$P[f\|H]$	-	prior probability
$P[D\|H]$	-	normalisation (independent of f)

The most probable image will be the $f(\omega)$ which satisfies

$$\frac{\delta P[f|DH]}{\delta f} = 0. \tag{2.3}$$

2.2 The likelihood function

For a large number of Monte Carlo measurements of a correlation function, the input data, D is expected to obey the Gaussian distribution law according to the central limit theorem.

$$P[D|fH] = \frac{1}{Z_L} e^{-L}, \tag{2.4}$$

$$L = \frac{1}{2} \sum_{i,j} (D(\tau_i) - F(\tau_i)) C_{ij}^{-1} (D(\tau_j) - F(\tau_j)), \tag{2.5}$$

where the indices i and j run over the actual time window used in the analysis, $\frac{\tau_{min}}{a} \leq i, j, \leq \frac{\tau_{max}}{a}$ (a is the lattice spacing). Define the number of data points used (for later use)

$$N_\tau = \frac{\tau_{max}}{a} - \frac{\tau_{min}}{a} + 1. \tag{2.6}$$

$D(\tau_i)$ is the averaged lattice data at time τ_i

$$D(\tau_i) = \frac{1}{N_{cfg}} \sum_{m=1}^{N_{cfg}} D^m(\tau_i), \tag{2.7}$$

where N_{cfg} is the total number of gauge configurations available and $D^m(\tau_i)$ is the data for the m^{th} configuration. $F(\tau_i)$ is the correlation function data calculated from f,

$$F(\tau) = \int_0^\infty K(\tau, \omega) f(\omega) d\omega, \tag{2.8}$$

$K(\tau, \omega)$ is the lattice kernel,

$$K(\tau_i, \omega) = e^{-\omega\tau}. \tag{2.9}$$

C_{ij} is the standard $N_\tau \times N_\tau$ covariance matrix defined as

$$C_{ij} = \frac{1}{N_{cfg}(N_{cfg} - 1)} \sum_{m=1}^{N_{cfg}} (D^m(\tau_i) - D(\tau_i))(D^m(\tau_j - D(\tau_j)). \tag{2.10}$$

For real lattice data it is essential to take account of the off-diagonal elements of C_{ij} as correlations between different τ are generally strong.

2.2.1 Normalisation

Z_L is a normalisation factor calculated from the integration of $P[D|fH]$ over D using the measure $[dD]$ given below.

$$[dD] \equiv \prod_{i=\frac{\tau_{min}}{a}}^{\frac{\tau_{max}}{a}} \tag{2.11}$$

$$Z_L = (2\pi)^{\frac{N_\tau}{2}} \sqrt{\det C} \tag{2.12}$$

2.2.2 Comparison to standard χ^2-fitting

If the prior probability, $P[f|H]$, is constant then maximising $P[f|DH]$ is identical to maximising $P[D|fH]$. This corresponds to minimising the likelihood, L, defined in Eq. 2.4. This is just standard χ^2-fitting. Generally the number of data points on a lattice, N_τ, is $\mathcal{O}(10)$ and too get good resolution on sharp features in f the number of points required is $\mathcal{O}(10^3)$. Hence χ^2-fitting alone will not work as there are many different f's which will minismise χ^2. The role of the non-constant prior probability is essential for MEM to overcome this ill-posed problem.

2.3 The prior probability

The prior probability used is defined as

$$P[f|H\alpha m] = \frac{1}{Z_S} e^{\alpha S}, \tag{2.13}$$

where α and m are auxiliary parameters which will be discussed in more detail later. S is the Shannon Jaynes entropy.

$$S = \int_0^\infty \left[f(\omega) - m(\omega) - f(\omega) \log \left(\frac{f(\omega)}{m(\omega)} \right) \right] d\omega. \tag{2.14}$$

On the lattice the discretised form of this is required

$$S = \sum_{l=1}^{N_\omega} \left[f_l - m_l - f_l \log \left(\frac{f_l}{m_l} \right) \right] \tag{2.15}$$

α is a real positive parameter and $m(\omega)$ is a real positive function known as the *default model* (sometimes called the *prior estimate*). α and m are part of the hypothesis, H, but are now written in explicitly to all terms involving H including the likelihood even though it is independent of both α and m. At the end of the calculation α will be integrated out and therefore eliminated from the final result. The default model, m, remains in the final result, but the sensitivity of the results against a change in m will be studied 3.2.3.

2.3.1 Discretisation of frequency, ω

As seen in Eq. 2.15, then the frequency ω is discretised as follows. There are N_ω points with equal spacing $\Delta\omega$ so $f_l \equiv f(\omega_l)$, $m_l \equiv m(\omega_l)$ and $\omega_l \equiv l\Delta\omega$.

2.3.2 Normalisation

Once again there is a normalisation factor, Z_S, involved in the definition of the prior probability (Eq. 2.13). The integration of $P[f|H\alpha m]$ over f using the measure $[df]$

(defined below) is normalised to unity.

$$[df] \equiv \prod_{l=1}^{N_\omega} \frac{df_l}{\sqrt{f_l}}, \tag{2.16}$$

$$Z_S \cong \left(\frac{2\pi}{\alpha}\right)^{\frac{N_\omega}{2}}. \tag{2.17}$$

In Appendix A two derivations for the Shannon-Jaynes entropy are given. Firstly by using the so-called "*monkey argument*" [36, 37, 38] and secondly a derivation based on an axiomatic argument [39].

2.4 MEM algorithm

There are three steps involved in calculating the image, $f(\omega)$.

2.4.1 Step 1 - Find the most probable $f(\omega)$ for a given α

This involves maximising

$$P[f|DH\alpha m] \propto \frac{1}{Z_S Z_L} e^Q, \tag{2.18}$$

$$Q \equiv \alpha S - L \tag{2.19}$$

Eq. 2.18 is obtained by combining Eqs. 2.2, 2.4 and 2.13 and noting that $P(D|H)$ is independent of f and is therefore an independent constant. It is easy to see from Eq. 2.19 that α just plays the role of the relative weight between S (which fits to $m(\omega)$) and L (which fits to D).

The most probable image for a given α is f_α, which satisfies the condition

$$\left.\frac{\delta Q}{\delta f(\omega)}\right|_{f=f_\alpha} = 0 \tag{2.20}$$

For a detailed description on the algorithm used for solving Eq. 2.20 see Sec. 2.5.

2.4.2 Step 2 - Calculate final image

Once f_α has been found the choice for the final result depends on the MEM procedure being used. The three options are

1. Classical - choose α such that $\chi^2 = N_\tau$ [36]
2. Historic - choose the f_α which maximises $P[\alpha|DHm]$ [36]
3. Bryan's - Perform an average over f and α weighted by $P(\alpha|DHm)$ [35]

If $P[\alpha|DHm]$ is fairly strongly dependent on α then Bryan's algorithm is required, and is used throughout this work.

Using Bryan's algorithm the final image, f_{out}, is defined as

$$f_{out} = \int [df] \int d\alpha f(\omega) P[f|DH\alpha m] P[\alpha|DHm], \tag{2.21}$$

$$\simeq \int d\alpha f_\alpha(\omega) P[\alpha|DHm]. \tag{2.22}$$

The last expression here is obtained by using the assumption that $P[f|DH\alpha m]$ is sharply peaked around $f_\alpha(\omega)$, which should be satisfied for good data with small errors. $P[\alpha|DHm]$ can be calculated using Bayes' theorem

$$P[\alpha|DHm] = \int [df] \frac{P[D|fH\alpha m]P[f|H\alpha m]P[\alpha|Hm]}{P[D|Hm]}, \tag{2.23}$$

$$\propto P[\alpha|Hm] \int [df] \frac{1}{Z_S Z_l} e^{Q}, \tag{2.24}$$

$$\propto P[\alpha|Hm] \exp\left[\frac{1}{2}\sum_k \log\frac{\alpha}{\alpha+\lambda_k} + \alpha S(f_\alpha) - L(f_\alpha)\right], \tag{2.25}$$

where the λ_k's are eigenvalues of a real symmetric matrix defined by

$$\Lambda_{l,l'} \equiv \sqrt{f_l}\frac{\partial^2 L}{\partial f_l \partial f_{l'}}\sqrt{f_l'}\Big|_{f=f_\alpha}. \tag{2.26}$$

Choosing prior probability for α

$P[\alpha|Hm]$ is known as the prior probability for α. There are two rules for choosing this, either the Laplace rule ($P[\alpha|Hm] = const$) or Jeffrey's rule ($P[\alpha|Hm] = \frac{1}{\alpha}$) [34]. This choice is arbitrary as long as $P[\alpha|DHm]$ is concentrated around its maximum $\hat{\alpha}$

2.4.3 Step 3 - Error analysis

An advantage of using MEM to construct the image $f(\omega)$ is that it allows one to analyse the statistical significance of the peaks found. Since neighbouring points are heavily correlated, the error needs to be calculated over some interval in ω [25]. First define the unweighted average of $f(\omega)$ over a region I in ω.

$$\langle f_\alpha \rangle_I \equiv \frac{\int [df] \int_I d\omega f(\omega) P[f|DH\alpha m]}{\int_I d\omega} \simeq \frac{\int_I d\omega f_\alpha(\omega)}{\int_I d\omega}, \tag{2.27}$$

where to get the last expression it has been assumed that $P[f|DH\alpha m]$ is strongly peaked around $f_\alpha(\omega)$.

The variance of $\langle f_\alpha \rangle_I$ is similarly estimated by

$$\langle (\delta f_\alpha)^2 \rangle_I \equiv \frac{\int [df] \int_{I\times I} \delta\omega\delta\omega' \delta f(\omega)\delta f(\omega') P[f|DH\alpha m]}{\int_{I\times I} \delta\omega\delta\omega'} \tag{2.28}$$

$$\simeq -\frac{\int_{I\times I} \delta\omega\delta\omega' \left(\frac{\delta^2 Q}{\delta f(\omega)\delta f(\omega')}\right)^{-1}_{f=f_\alpha}}{\int_{I\times I} \delta\omega\delta\omega'}, \tag{2.29}$$

where $\delta f(\omega) = f(\omega) - f_\alpha(\omega)$. To get the last expression, the Gaussian approximation has been made for $P[f|DH\alpha m]$. As the final output image f_{out} is an average over α, the same is done to find the final value of the variance

$$\langle (\delta f_{out})^2 \rangle_I \equiv \int d\alpha \langle (\delta f_\alpha)^2 \rangle_I P[\alpha|DHm]. \tag{2.30}$$

These errors are shown on the plot of the spectral function as an ordinary error bar cross, but the meaning is slightly different. The horizontal position shows the central position of the peak, and the extent shows the region I which was averaged over (chosen to be the full width at half maximum). The vertical position indicates the average height of the peak (see Eq. 2.27) and the extent shows the variance (see Eq. 2.30). The way in which these errors are interpreted is as follows. If the variance of the peak is much smaller than the average height then the peak is a physical feature, but if it is larger then the peak is considered to be statistically insignificant.

2.4.4 Condition for integrating over α

The range of α over which the averaging is performed is determined using the criterion

$$P[\alpha|DHm] \geq 1\% \times P[\hat{\alpha}|DHm]. \tag{2.31}$$

This probability is then renormalised so that

$$\int_{\alpha_{min}}^{\alpha_{max}} d\alpha P[\alpha|DHm] = 1, \tag{2.32}$$

then the integration in Eq. 2.22 is carried out.

2.5 Maximising Q

This is the most computationally intensive part of the algorithm since the functional space of $f(\omega)$, in which the global maximum of Q is found, has typically $\mathcal{O}(10^3)$ degrees of freedom. Fortunately Bryan [35] found that, by using a singular value decomposition (SVD) on the kernel, K, the search space can be restricted to, at most, the number of data points $\sim \mathcal{O}(10)$. In this subsection, the algorithm originally proposed by Bryan [35] is followed.

2.5.1 Extremum condition

The extremum condition again is

$$\frac{\delta Q}{\delta f_l} = 0, \tag{2.33}$$

$$\text{i.e.} \qquad \alpha \frac{\partial S}{\partial f_l} - \frac{\partial L}{\partial f_l} = 0. \tag{2.34}$$

Using

$$F(\tau, \mathbf{k}) \equiv \int_0^\infty K(\tau, \omega) f(\omega, \mathbf{k}) d\omega \tag{2.35}$$

where $\mathbf{k}$ is the momentum of the channel being studied. Eq. 2.34 can be re-written as

$$\alpha \log\left(\frac{f_l}{m_l}\right) - \sum_{i=\frac{\tau_{min}}{a}}^{\frac{\tau_{max}}{a}} K_{il} \frac{\partial L}{\partial F_i} = 0. \tag{2.36}$$

2.5.2 SVD of lattice kernel

The explicit form of the SVD is shown below.

$$\begin{aligned} K^T &= UWV^T \\ &= \begin{pmatrix} u_{11} & \dots & u_{1N_\tau} \\ \vdots & \ddots & \vdots \\ u_{N_\omega 1} & \dots & u_{N_\omega N_\tau} \end{pmatrix} \begin{pmatrix} w_1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & w_{N_\tau} \end{pmatrix} \begin{pmatrix} v_{11} & \dots & v_{N\tau 1} \\ \vdots & \ddots & \vdots \\ v_{1N_\tau} & \dots & v_{N_\tau N_\tau} \end{pmatrix}, \end{aligned} \tag{2.37}$$

where V and U are $N_\omega \times N_\tau$ and $N_\tau \times N_\tau$ orthogonal matrices respectively (N_τ is the number of time slices defined in Eq. 2.6 and N_ω is the number of points in the spectral function, $f(\omega)$). W is an $N_\tau \times N_\tau$ diagonal matrix with diagonal elements $W_{ii} = w_i, i = 1, \dots, N_\tau$. The w_i are conventionally ordered $w_1 \geq w_2 \geq \dots \geq w_{N_\tau}$. For a matrix close to singularity only $s \leq N_\tau$ diagonal elements will be non-zero. These are known as the singular values of K [40] and $s = rank(K)$ and this s-dimensional

space spanned by the first s columns of U will be referred to as the *"singular space"* for convenience.

2.5.3 Bryan's algorithm

The solution to Eq. 2.36 can be represented in terms of a new variable $\mathbf{u}$ where $\log\left(\frac{f}{m}\right) = K^T\mathbf{u}$. From

$$\nabla L = \frac{\partial \mathbf{F}}{\partial f}\frac{\partial L(\mathbf{F},\mathbf{D})}{\partial \mathbf{F}} = K^T\frac{\partial L(\mathbf{F},\mathbf{D})}{\partial \mathbf{F}}, \tag{2.38}$$

it is obvious that ∇L lies in the singular space of K and thus the problem can be reduced into s-dimensions. So

$$f_i = m_i \exp\sum_{t=1}^{s} V_{it}u_t. \tag{2.39}$$

Now writing ∇S in terms of $V\mathbf{u}$ and then using the fact that the columns of V are orthonormal Eq. 2.36 becomes

$$-\alpha V\mathbf{u} = VWU^T\frac{\partial L(\mathbf{F},\mathbf{D})}{\partial \mathbf{F}}, \tag{2.40}$$

$$-\alpha\mathbf{u} = WU^T\frac{\partial L(\mathbf{F},\mathbf{D})}{\partial \mathbf{F}} = WU^TC_{ij}^{-1}\underbrace{(\overbrace{Kf}^{\mathbf{F}} - \mathbf{D})}_{\frac{\partial L(\mathbf{F},\mathbf{D})}{\partial \mathbf{F}}} = \mathbf{g}. \tag{2.41}$$

2.5.4 Newton search for solution

To solve Eq. 2.41 by fixed-point iteration (i.e. starting with a trial $\mathbf{u}$, calculating $\mathbf{g}$ and then calculating an improved $\mathbf{u}$ via $\mathbf{u} = \frac{\mathbf{g}}{\alpha}$) will not work since the convergence criterion is $\left|\frac{\partial \mathbf{g}}{\partial \mathbf{u}}\right|_{\text{largest singular value}} < \alpha$.

However a Newton search method can be used to solve Eq. 2.41. The increment at each iteration is given by

$$J\delta\mathbf{u} = -\alpha\mathbf{u} - \mathbf{g}, \tag{2.42}$$

where J is the Jacobian of the system and is given by

$$J = \alpha I + \frac{\partial \mathbf{g}}{\partial \mathbf{u}} \tag{2.43}$$

$$\frac{\partial \mathbf{g}}{\partial \mathbf{u}} = WU^T C_{ij}^{-1} \frac{\partial \mathbf{F}}{\partial \mathbf{f}} \frac{\partial \mathbf{f}}{\partial \mathbf{u}}, \tag{2.44}$$

and by the chain rule

$$\frac{\partial \mathbf{f}}{\partial \mathbf{u}} = diag\{f\}V. \tag{2.45}$$

So

$$\frac{\partial \mathbf{g}}{\partial \mathbf{u}} = WU^T C_{ij}^{-1} \underbrace{UWV^T}_{\frac{\partial \mathbf{F}}{\partial \mathbf{f}} = K} diag\{f\}V = MT, \tag{2.46}$$

where the $s \times s$ matrices M and T are defined as

$$M = WU^T C_{ij}^{-1} UW \tag{2.47}$$

$$T = V^T diag\{f\} V. \tag{2.48}$$

The equation to solve is now in the form

$$(\alpha I + MT)\delta\mathbf{u} = -\alpha\mathbf{u} - \mathbf{g}. \tag{2.49}$$

To get Eq. 2.49 a second order approximation has been used in the Newton search. To ensure that this approximation remains true, the size of the increment $\delta\mathbf{u}$ needs to be restricted. This is achieved by augmenting J with multiples of the identity matrix [41]

$$((\alpha + \mu)I + MT)\delta\mathbf{u} = -\alpha\mathbf{u} - \mathbf{g}, \tag{2.50}$$

where μ, known as a Marquardt–Levenburg parameter, is chosen such that $\delta\mathbf{u}T\delta\mathbf{u} \leq \mathcal{O}(\sum\mathbf{f})$ (i.e. the increment size $\delta\mathbf{u}$ is small enough to guarantee that the lowest order approximation in the Newton search method is valid).

2.5.5 Diagonalising Eq. 2.50

The search for $\delta\mathbf{u}$ can be optimised by diagonalising Eq. 2.50 so that only $\mathcal{O}(s)$ operations are required for each α - μ pair tried rather than $\mathcal{O}(s^3)$. This is done by first diagonalising T by solving the eigenvalue problem

$$TP = P\Xi, \tag{2.51}$$

where $\Xi = diag\{\xi\}$ and $P^TP = I$. Next define B

$$B = diag\{\xi^{1/2}\}P^TMPdiag\{\xi^{1/2}\}, \tag{2.52}$$

and solve the further eigenproblem

$$BR = R\Lambda, \tag{2.53}$$

where $\Lambda = diag\{\lambda\}$ and $R^TR = I$. Next define the following quantity

$$Y = Pdiag\{\xi^{-1/2}\}R, \tag{2.54}$$

and note that the following relationships hold

$$T = Y^{-T}Y^{-1} \tag{2.55}$$

$$\Lambda = Y^{-1}MY^{-T}. \tag{2.56}$$

The diagonalised form of Eq. 2.50 is

$$((\alpha + \mu)I + \Lambda)Y^{-1}\delta\mathbf{u} = -\alpha Y^{-1}\mathbf{u} - Y^{-1}\mathbf{g}. \tag{2.57}$$

This gives s independent equations for the components of $Y^{-1}\delta\mathbf{u}$ and the step length is now given by

$$\delta\mathbf{u}^T T\delta\mathbf{u} = \delta\mathbf{u}^T Y^{-T}Y^{-1}\delta\mathbf{u} = \left|Y^{-1}\delta\mathbf{u}\right|^2 \tag{2.58}$$

2.5.6 Getting $\delta\mathbf{u}$ from $Y^{-1}\delta\mathbf{u}$

Often, in practice, $\mathbf{f}$ has a high dynamic range. In our case $\mathbf{f}$ is a mesonic spectral function, and is therefore likely to have a high peak for the ground state. In such a situation, T (defined in Eq. 2.48) may be close to singular, which means that some of the ξ calculated in Eq. 2.51 are effectively zero. In such cases merely multiplying the answer for $Y^{-1}\delta\mathbf{u}$ by Y would lead to a numerical instability as Y contains a $\xi^{-1/2}$ term. The solution to this problem is to notice that a simple rearrangement of Eq. 2.50 gives

$$\begin{aligned}(\alpha+\mu)\delta\mathbf{u} &= -\alpha\mathbf{u} - \mathbf{g} - MT\delta\mathbf{u} \\ &= -\alpha\mathbf{u} - \mathbf{g} - MY^{-T}[Y^{-1}\delta\mathbf{u}].\end{aligned} \tag{2.59}$$

2.5.7 How to deal with small/negative ξ's

When there are very small (or even negative) values for ξ they should be considered to be zero. In such a situation the matrix of eigenvectors from T, P should be partitioned into 2 parts. One part for those associated with the zero ξ's (the null

space, designated with a subscript 0) and another part for those associated with the non-zero ξ's (the non-null space, designated with a subscript 1). i.e.

$$\begin{aligned} T &= P\Xi P^T \\ &= \begin{pmatrix} P_1 & P_0 \end{pmatrix} \begin{pmatrix} \Xi_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} P_1^T \\ P_0^T \end{pmatrix}. \end{aligned} \tag{2.60}$$

P is partitioned in this way (opposite to that given in Bryan [35]) since the eigenvalues are organised largest to smallest by the eigenvalue subroutine used [41].

Now multiplying Eq. 2.50 on the left by P^T, writing $\mathbf{p} = P^T\mathbf{u}$ and using Eq. 2.60 to partition gives.

$$(\alpha+\mu)\begin{pmatrix} \delta\mathbf{p}_1 \\ \delta\mathbf{p}_0 \end{pmatrix} + \begin{pmatrix} P_1^T \\ P_0^T \end{pmatrix} M \begin{pmatrix} P_1 & P_0 \end{pmatrix} \begin{pmatrix} \Xi_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \delta\mathbf{p}_1 \\ \delta\mathbf{p}_0 \end{pmatrix} = -\alpha\mathbf{p} - P^T\mathbf{g}. \tag{2.61}$$

This gives two equations, Eq. 2.62 for $\delta\mathbf{p}_1$ which is of the same form as Eq. 2.50 and Eq. 2.63 into which the answer for $\delta\mathbf{p}_1$ is substituted to obtain the answer for $\delta\mathbf{p}_0$.

$$(\alpha+\mu)\delta\mathbf{p}_1 + P_1^T M P_1 \Xi_1 \delta\mathbf{p}_1 = -\alpha\mathbf{p}_1 - P_1^T\mathbf{g} \tag{2.62}$$

$$(\alpha+\mu)\delta\mathbf{p}_0 + P_0^T M P_1 \Xi_1 \delta\mathbf{p}_1 = -\alpha\mathbf{p}_0 - P_0^T\mathbf{g} \tag{2.63}$$

Calculating the right hand side of these two equations can be done in one step since there is no mixing between the null and non-null space.

$$
\begin{aligned}
-\alpha\mathbf{p} - P^T\mathbf{g} = P^T(-\alpha\mathbf{u} - \mathbf{g}) &= \begin{pmatrix} P_1^T \\ P_0^T \end{pmatrix}(-\alpha\mathbf{u} - \mathbf{g}) \\
&= \begin{pmatrix} P_1^T(-\alpha\mathbf{u} - \mathbf{g}) \\ P_0^T(-\alpha\mathbf{u} - \mathbf{g}) \end{pmatrix} \\
&= \begin{pmatrix} -\alpha\mathbf{p}_1 - P_1^T\mathbf{g} \\ -\alpha\mathbf{p}_0 - P_0^T\mathbf{g} \end{pmatrix} \qquad (2.64)
\end{aligned}
$$

For similar reasons the product of P^TMP can also be calculated in one step to give both $P_1^TMP_1$ and $P_0^TMP_1$

$$
\begin{aligned}
P^TMP &= \begin{pmatrix} P_1^T \\ P_0^T \end{pmatrix}(M)\begin{pmatrix} P_1 & P_0 \end{pmatrix} \\
&= \begin{pmatrix} P_1^T \\ P_0^T \end{pmatrix}\begin{pmatrix} MP_1 & MP_0 \end{pmatrix} \\
&= \begin{pmatrix} P_1^TMP_1 & P_1^TMP_0 \\ P_0^TMP_1 & P_0^TMP_0 \end{pmatrix}. \qquad (2.65)
\end{aligned}
$$

2.5.8 Diagonalising Eq. 2.62

This follows a very similar procedure to that used in Sec. 2.5.5.

Firstly, diagonalise the matrix T as in Eq. 2.51 and define B as

$$B = diag\{|\xi^{1/2}|\}P^TMPdiag\{|\xi^{1/2}|\}. \qquad (2.66)$$

Diagonalise B only in the non-null space

$$B_1 R_1 = R_1 \Lambda_1. \tag{2.67}$$

Now define Y as

$$Y_1 = diag\{\xi^{-1/2}\}_1 R_1, \tag{2.68}$$

this differs from Eq. 2.54 by a factor of P since $\delta\mathbf{u}$ has already been multiplied by this. Once again note the relationships

$$\Xi_1 = Y_1^{-T} Y_1^{-1} \tag{2.69}$$

$$\Lambda_1 = Y_1^{-1} P_1^T M P_1 Y_1^{-T} \tag{2.70}$$

Hence the diagonal form of Eq. 2.62 is

$$((\alpha + \mu)I + \Lambda) Y_1^{-1} \delta\mathbf{p}_1 = -\alpha Y_1^{-1} \mathbf{p}_1 - Y_1^{-1} P_1^T \mathbf{g}. \tag{2.71}$$

Once again the same trick is used to get from $Y_1^{-1}\delta\mathbf{p}_1$ to $\delta\mathbf{p}_1$

$$(\alpha + \mu)\delta\mathbf{p}_1 = -\alpha\mathbf{p}_1 - P_1^T\mathbf{g} - P_1^T M P_1 Y_1^{-T}[Y_1^{-1}\delta\mathbf{p}_1]. \tag{2.72}$$

The answer for $\delta\mathbf{p}_0$ is obtained by substituting δp_1 into Eq. 2.63 and $\delta\mathbf{u}$ is then constructed by left multiplying by P.

2.6 Uniqueness of the solution in MEM

In order to prove the uniqueness of the MEM solution, the following proposition must be proved

Proposition:

Consider a real smooth function F with n real variables, $F(x_1, \ldots, x_n) \in \mathbb{R}$ with $(x_1, \ldots, x_n) \in \mathbb{R}^n$. Suppose that for any $y_i \in \mathbb{R}$

$$\sum_{i,j=1}^{n} y_i \frac{\partial^2 F}{\partial x_i \partial x_j} y_j < 0, \tag{2.73}$$

then the function F only has one maximum if it exists, i.e. the solution of

$$\frac{\partial F}{\partial x_i} = 0 \quad (i = 1, \ldots, n), \tag{2.74}$$

is unique if it exists.

Proof:

Assume that there are at least two solutions for Eq. 2.74. Use any two solutions, $\vec{x}_1$ and $\vec{x}_2$, to define an interpolation

$$x(t) = \vec{x}_1 + t(\vec{x}_2 - \vec{x}_1) \equiv \vec{x}_1 + t\vec{y}, \tag{2.75}$$

and $G(t) \equiv F(\vec{x}(t))$. Using the assumption that $dG(t)/dt$ is continuous, differentiable in $[0, 1]$ and satisfies

$$\left.\frac{dG}{dt}\right|_t = 0 = \left.\frac{dG}{dt}\right|_{t=1} = 0, \tag{2.76}$$

Rolle's theorem states that there exists at least one $t \in [0,1]$ such that

$$\frac{d^2G(t)}{dt^2} = \sum_{i,j=1}^{n} y_i \left.\frac{\partial^2 F}{\partial x_i \partial x_j}\right|_{\vec{x}=\vec{x}(t)} y_j = 0. \tag{2.77}$$

Comparing Eqs. 2.73 and 2.77 there is a contradiction, hence there cannot be two or more solutions to Eq. 2.74. So if there is a solution then it is the global maximum of F

Now apply this to the search for the global maximum of $Q = \alpha S - L$ to prove its uniqueness. For an arbitrary N_ω–dimensional non-zero real vector $\vec{z} = (z_1, \ldots, z_{N_\omega})$, αS satisfies (see Eq. 2.15)

$$\sum_{l,l'=1}^{N_\omega} z_l \frac{\partial^2(\alpha S)}{\partial f_l \partial f_{l'}} z_{l'} = -\alpha \sum_{l=1}^{N_\omega} \frac{z_l^2}{f_l} < 0, \tag{2.78}$$

where we have used $0 \leq f_l < \infty$ and $0 < \alpha < \infty$. Importantly, notice that the left hand side of Eq. 2.78 never becomes zero. From the definition of the likelihood L in Eq. 2.4

$$\sum_{l,l'=1}^{N_\omega} z_l \frac{\partial^2(-L)}{\partial f_l \partial f_{l'}} z_{l'} = -\sum_{i=1}^{N_\tau} \frac{\bar{z}_i^2}{\bar{\sigma}_i} \leq 0, \quad \text{with} \quad \bar{z}_i = \sum_{l=1}^{N_\omega} \bar{K}_{il} z_l \tag{2.79}$$

The left hand side of Eq. 2.79 is zero in the direction $\bar{z}_i = 0(i = 1, \ldots, N_\tau)$, of which there are many since the rank of $\bar{K}$ is at most N_τ, much smaller than the N_ω–dimension of z_l. Thus $-L$ has many flat directions and there is no unique maximum for $-L$ as a function of f_l. However, considering both Eqs. 2.78 and 2.79 together, the maximum for Q is unique, if it exists, due to the proposition just proved.

Chapter 3

Testing the Maximum Entropy Method

In this chapter the ability of the MEM algorithm to reproduce the correct image, and the dependence it has on the quality of the input data, will be analysed. This will involve two sets of input spectral functions:

1. Delta function
2. QCD-like spectral function

The covariance matrix, C_{ij}, will, for simplicity, be assumed to be diagonal. It is important to remember that the off-diagonal elements of C_{ij} play an important role when analysing real lattice QCD data.

The assumptions made in the derivation of the MEM algorithm, i.e. the dependence on the default model (§ 3.2.3) and prior probability for α (§ 3.2.4) will be tested. The output image will be analysed and quantities such as the mass and width of the peaks will be extracted.

3.1 Test data generation

3.1.1 Delta function

Outlined below is the manner in which the test data, based on a delta-function-like spectral function, was generated. The aim of this data is to model correlators, $D(\tau)$, consisting of a discrete set of states, which can be expressed as

$$D(\tau_i) = \sum_{j=1}^{N_{\text{states}}} \frac{Z_j}{2M_j} e^{-M_j \tau_i} \tag{3.1}$$

where $\tau_i = \{0, 1, \ldots, N_\tau\}$. Two cases will be considered (i) a single pole at $M_1 = 0.5$ and (ii) double poles at $M_1 = 0.5$ and $M_2 = 1.0$.

1. In each case random noise is generated on all input Z_i and M_i giving N_{cfg} copies of Z_i and M_i.

2. These are then used to calculate $D^k_{\text{in}}(\tau)$ from Eq. 3.1 for $k = 1, \ldots, N_{\text{cfg}}$.

3. The central value is then calculated as

$$\bar{D}_{\text{in}}(\tau_i) = \sum_{k=1}^{N_{\text{cfg}}} D^k_{\text{in}}(\tau_i) \tag{3.2}$$

Note: Standard χ^2-fitting should work almost perfectly for this style of data, so the results obtained for Z and M will be compared against those obtained from standard exponential fits. Analysis will also be performed on the width of the peak to test whether the width in this case is purely statistical as expected.

3.1.2 QCD-like data

The second set of data was generated using the procedure outlined below.

1. Starting with an input image $f_{\rm in}(\omega) = \omega^2 \rho_{\rm in}(\omega)$. (The factor of ω^2 is expected from the dimension of mesonic operators). Test data is then calculated from this spectral function via

$$D_{\rm in}(\tau_i) = \int_0^{\omega_{\rm max}} K(\tau_i, \omega) f_{\rm in}(\omega) d\omega, \tag{3.3}$$

   where the lattice kernel is defined as

$$K(\tau_i, \omega) = e^{-\omega \tau_i} \tag{3.4}$$

   and $\omega_{\rm max}$ is taken to be large enough so that $\rho_{\rm in}(\omega)$ does not show appreciable variation. The detailed form for $\rho_{\rm in}(\omega)$ will be given in Eq. 3.7.

2. Gaussian noise with variance $\sigma(\tau_i)$ (defined below) is then added to each $D_{\rm in}(\tau_i)$ to create the test data.

$$\sigma(\tau_i) = b \times D_{\rm in}(\tau_i) \times \frac{\tau_i}{\Delta\tau}, \tag{3.5}$$

   where the dependence on τ_i is introduced in order to incorporate the fact that the error in lattice correlation functions increases with τ.

3. Since this data is generated with an input image f_{in}, we can compare directly the output image, f_{out} with this. To give a numerical accuracy on f_{out} we define a "χ^2" between input and output images

$$r \equiv \int_0^{\omega_{\max}} [\rho_{\text{out}}(\omega) - \rho_{\text{in}}(\omega)]^2 d\omega. \tag{3.6}$$

As in [42] the input spectral function used, $\rho_{in}(\omega)$, is from the vector channel of the e^+e^- annihilation and is defined as follows:

$$\rho_{in}(\omega) = \frac{2}{\pi}\left[F_\rho^2 \frac{\Gamma_\rho m_\rho}{(\omega^2 - m_\rho^2)^2 + \Gamma_\rho^2 m_\rho^2} + \frac{1}{8\pi}\left(1 + \frac{\alpha_s}{\pi}\right)\frac{1}{1 + e^{(\omega_0 - \omega)/\delta}}\right], \tag{3.7}$$

where F_ρ is the residue of the ρ meson resonance defined by

$$\langle 0|\bar{d}\gamma_\mu u|\rho\rangle = \sqrt{2}F_\rho m_\rho \epsilon_\mu \equiv \sqrt{2} f_\rho m_\rho^2 \epsilon_\mu \tag{3.8}$$

with the polarisation vector ϵ_μ. Γ_ρ includes a θ function which represents the threshold of $\rho \to \pi\pi$ decay

$$\Gamma_\rho(\omega) = \frac{g_{\rho\pi\pi}^2}{48\pi} m_\rho \left(1 + \frac{2m_\pi^2}{\omega^2}\right)^{\frac{3}{2}} \theta(\omega - 2m_\pi). \tag{3.9}$$

The empirical values of the parameters are

$$\begin{aligned} m_\rho &= 0.77 \text{ GeV}, \quad m_\pi = 0.14 \text{ GeV}, \\ g_{\rho\pi\pi} &= 5.45, \qquad \alpha_s = 0.3, \\ \omega_0 &= 1.3 \text{ GeV}, \qquad \delta = 0.2 \text{ GeV}, \end{aligned} \tag{3.10}$$

where for simplicity α_s has been assumed to be independent of ω.

The default model, m, used throughout these tests is taken to be of the form $m(\omega) = m_0\omega^2$ which is motivated by the asymptotic behaviour of $f_{\text{in}}(\omega)$. m_0 is taken to be the value obtained from the large ω limit

$$\lim_{\omega\to\infty} \rho_{in}(\omega) = \frac{2}{\pi}\frac{1}{8\pi}\left(1 + \frac{\alpha_s}{\pi}\right) = 0.277. \tag{3.11}$$

The frequency space is discretised with $\omega_{\max} = 6.0$, $N_\omega = 600$ and therefore $\Delta\omega = 0.01$.

3.2 The tests

3.2.1 Simple pole analysis

# configs	Mass (1-exp)	Mass (MEM)
100	0.504(2)	0.501(4)
150	0.507(2)	0.503(3)
200	0.508(2)	0.506(3)
250	0.509(2)	0.506(3)
300	0.508(2)	0.505(2)

Table 3.1: Comparison of MEM with single-exponential fit to data generated with a simple pole at $M_1 = 0.5$ for increasing number of configurations

In this section the delta function data will be used to assess MEM's ability to find simple poles in noisy correlation function data. It will be compared directly with single- and multi-exponential fits, which work extremely well on data of this kind. Table 3.1 compares the results obtained from single-exponential and MEM fits to data generated with a simple pole at M=0.5. The double pole comparison with a two-exponential fit is tabulated in Table 3.2. The value obtained for the masses from the two methods agrees within errors

# configs	Mass 1 (2-exp)	Mass 1 (MEM)	Mass 2 (2-exp)	Mass 2 (MEM)
100	0.491(1)	0.510(4)	1.003(12)	1.003(6)
150	0.491(1)	0.504(6)	0.998(10)	1.003(5)
200	0.492(1)	0.497(4)	1.001(9)	1.002(5)
250	0.491(1)	0.496(2)	1.000(8)	1.001(4)
300	0.491(1)	0.495(3)	1.003(8)	1.001(3)

Table 3.2: Comparison of MEM with two-exponential fit to data generated with simple poles at $M_1 = 0.5$ and $M_2 = 1.0$ for increasing number of configurations

3.2.2 Input data quality

The quality of the input data will be altered in two ways. Firstly the number of time slices generated which increases the number of input data used to reconstruct the spectral function. And secondly the noise level will be altered. This is the same as a change in the number of configurations, smaller noise levels equates to more configurations.

Fig. 3.1 shows the dependence of the output spectral function on the quality of the input data for the QCD-like data. As expected increasing N_τ and decreasing the noise both result in an improved image, but the most drastic improvements occur when the noise is decreased.

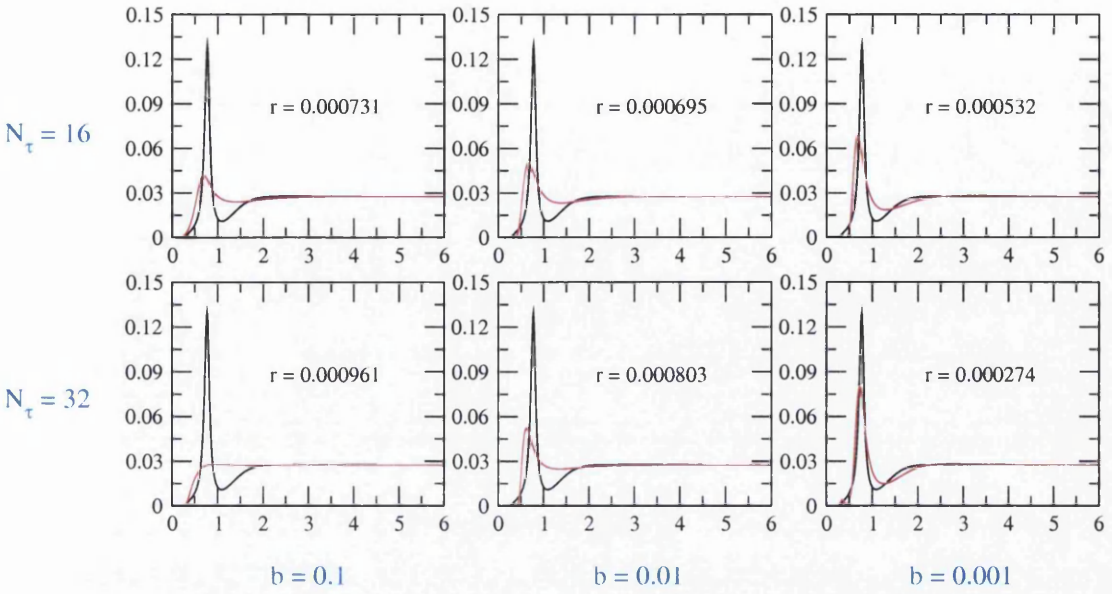

Figure 3.1: Input (black) and output (red) spectral function ($\rho(\omega)$) compared for different N_τ and noise level (b). r (defined in Eq. 3.6) is a measure of how close the two are.

3.2.3 Changing the default model

For these tests the QCD-like data will be used as it demonstrates most clearly the consequences of choosing the default model form badly. The functional form of the default model used in this analysis is $m(\omega) = m_0\omega^2$ motivated by the asymptotic form for $\rho_{in}(\omega)$.

Fig. 3.2 shows the output image obtained for 5 different values of the default model parameter, m_0. It is clear that the choosing the correct default model is important for the high energy regime, whereas the ground state peak is relatively stable under a change of default model. MEM attempts to set $\lim_{\omega\to 6.0}\rho(\omega) = m_0$ while keeping the total area under the spectral function constant. The result of this is that the continuum exhibits a ringing behaviour.

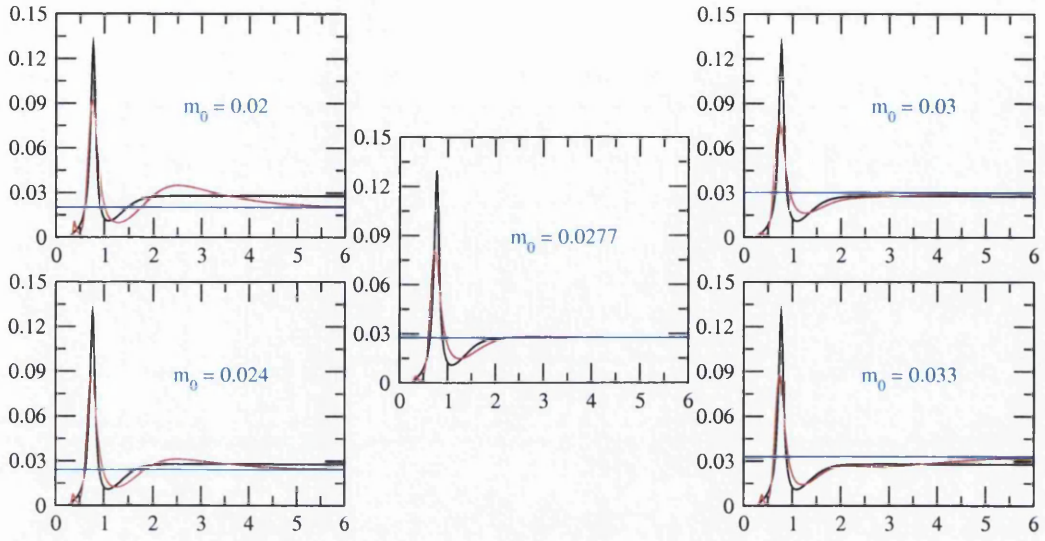

Figure 3.2: Input (black) and output (red) spectral function ($\rho(\omega)$) compared for different values of the default model parameter m_0. The blue horizontal line drawn is $\rho(\omega) = m_0$.

3.2.4 Changing the prior probability for α

Next the effects on the spectral function obtained when the prior probability for α is altered is studied. Two different forms for this probability can be used, Jeffrey's and Laplace (see § 2.4.2). Fig. 3.3 shows the comparison using these 2 definitions for the QCD-like test data. Even though the probability density is quite different (see inset graph) the resultant spectral function barely changes, in fact the 2 lie on top of each other and are indistinguishable by eye, using Eq. 3.6 to get a numerical value on the difference, $r = 2 \times 10^{-7}$. Thus the two images are almost identical making the choice of prior probability unimportant.

3.2.5 Spectral width

In this test the width of the peak obtained from the delta function data will be analysed for an increasing number of configurations. Since a delta function has zero

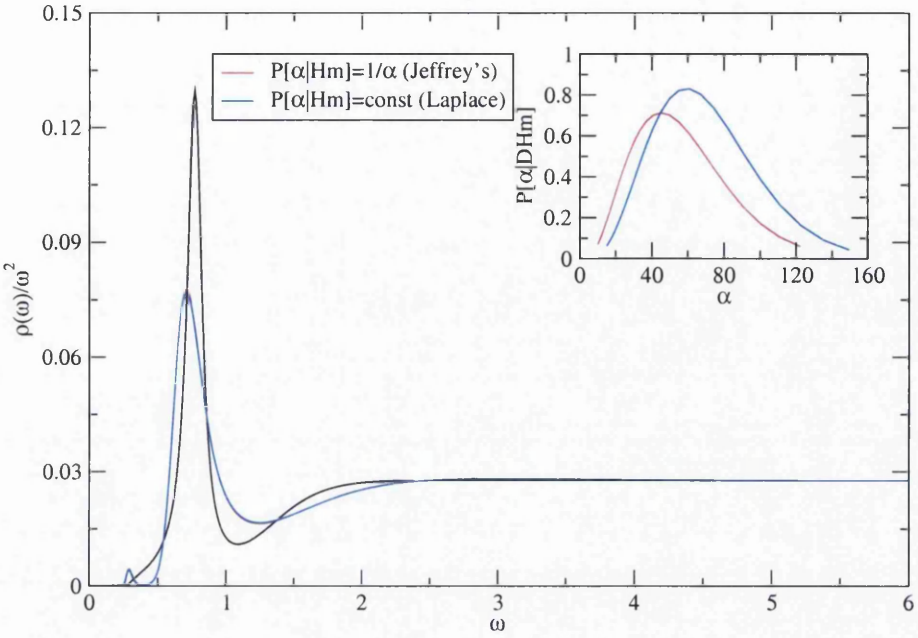

Figure 3.3: Input (black) and output using Jeffrey's (red) and the Laplace (blue) prior for $P[\alpha|Hm]$. Inset is the probability density used in the weighted average in each case

width an extrapolation to an infinite number of configurations will be performed to test whether the width is purely statistical. Fig. 3.4 is a plot of this extrapolation, the result of which is consistent with zero although the errors are large.

3.3 Summary

The following conclusions have been drawn from the tests performed on the MEM algorithm:

- MEM can find simple poles and the mass agrees within errors with standard single-exponential fits.
- Improving the quality of the data, both N_{cfg} and N_τ, results in better a quality image.

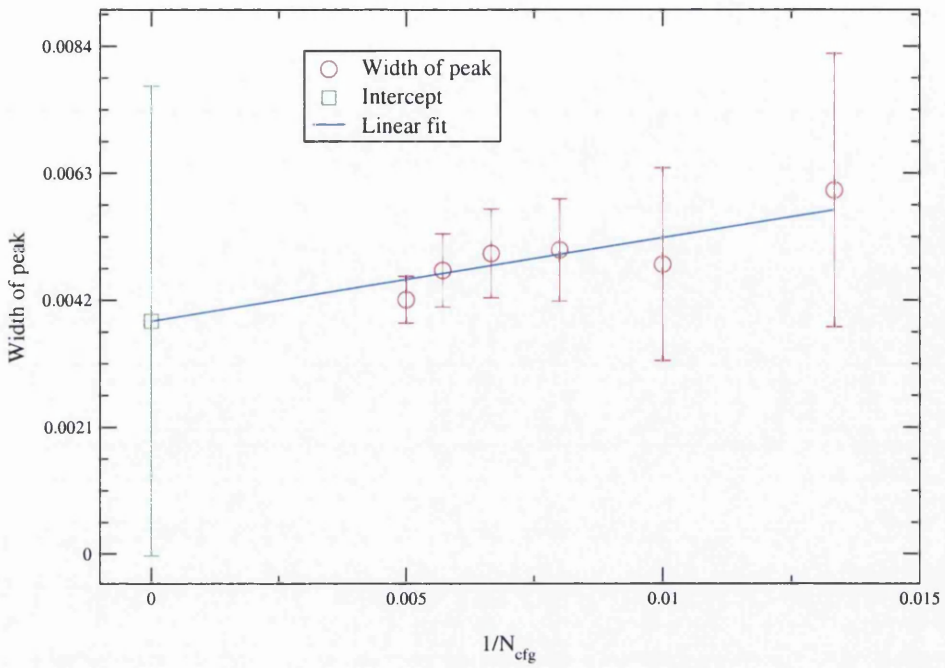

Figure 3.4: Extrapolation of the width in delta function data to an infinite number of configurations.

- Selecting the correct default model is important, any a priori knowledge about the model can and should be input here.
- Either form given here for the $P[\alpha|Dm]$ can be used with very little effect on the output image.
- When fitting to a simple pole, an extrapolation to infinite N_{cfg} agrees with zero, so the width seen is presumably due to the noisy data.

Chapter 4

Modelling QCD: The Four-Fermion Interaction

Due to the incredible complexity of QCD [1], studies of the finite temperature transition have been unable to yield quantitative claims for the universality class of the two light quark flavour transition. This is unfortunate since it is believed that the high temperature QCD phase transition has a number of interesting features [45, 46]. In addition to this, only very slow progress has been made in lattice simulations at finite chemical potential [47]. It is therefore advantageous to approach the problem of chiral symmetry restoration at finite temperature and density by using a "toy" model. Such studies should give a better understanding of the factors which might play a crucial role in the more complex gauge theory systems. In order to produce high quality data and study a large range of parameters this model has been simplified as much as possible. This model is interesting on its own, since it is non-trivial, non-asymptotically free and strongly interacting.

The four-fermion model appears frequently in physics. It was introduced as a model of β-decay [48] by Fermi. Bardeen *et al.* used it in condensed matter physics

[1] Note that there has been some very recent progress in this area [43, 44]

to model superconductivity [49]. Nambu and Jona-Lasinio applied many of the ideas in the BCS model of superconductivity to particle physics [50]. In particular, they proposed the four-fermion interaction could be used to explain the smallness of the pion mass, why the nucleon is so heavy and also to construct the meson as a particle–anti-particle bound state of some "originally massless fermions". In spite of the fact that it lacks confinement, the Nambu–Jona-Lasinio (NJL) model is believed to be an effective theory of quarks and gluons at intermediate energies [51, 52], i.e. between the asymptotic freedom and confinement regions, and it successfully describes certain aspects of hadron structure [53]. Four-fermion theories have appeared in recent discussions of dynamical mass generation in the Standard model, in such scenarios as walking technicolor [54, 55] and the top mode standard model, in which the Higgs scalar is a $t\bar{t}$ bound state [56, 57].

The NJL model is a trivial theory in four-dimensions [58], i.e. the renormalised coupling goes to zero in the continuum limit, hence it becomes an effective field theory only for scales $\ll \Lambda \approx 1GeV$, the UV scale. The low temperature regime is dominated by the lightest particles. If the restoration temperature is of the order of 100MeV, then the contribution of heavier particles like ρ mesons is exponentially suppressed. In that sense, the universal properties of chiral symmetry restoration in QCD could well be described by an effective theory like the NJL model [59, 60].

So why not simplify the model down to two-dimensions? In this case there are conceptual difficulties, e.g. in the Z_2 case the symmetry restoration is now dominated by the materialisation of kink–anti-kink states [61], which are composites of the fundamental fermion fields. At any non-zero temperature and for any finite number of fermion species N_f the chiral symmetry is restored due to the condensation of the

kinks [62, 63].

In three-dimensions, the four-fermion model is non-renormalisable in a "weak coupling expansion", but becomes renormalisable in the $1/N_f$ expansion [64]. Therefore, many quantities such as the fermion mass, composite particle masses and propagators, etc. can be calculated analytically. The basis of the $1/N_f$ expansion is the fact that the partition function of models like the four-fermion model can be expressed in the generic form [64],

$$Z = \int [d\sigma] \exp\left[\frac{-1}{\hbar} N_f S_{eff}[\sigma]\right], \tag{4.1}$$

by integrating out the fermions. The factor N_f in the exponential allows a saddle point approximation for the large N_f limit. The factor $1/\hbar$ also permits the same type of approximation since $\hbar$ is small. The expansion based on $\hbar$ is known as the "loop approximation". The counting of the order is different in these two expansions since the factor N_f can arise from different sources than the factor $1/\hbar$. For example, while a factor N_f can arise from a sum over flavours, the propagator contributes no factors of N_f. In fact, Feynman diagrams of the same order in $1/N_f$ can include diagrams of higher (up to infinite) order in $\hbar$. In this regard, the $1/N_f$ approximation is deemed to be a non-perturbative expansion in $\hbar$.

4.1 Essential properties

In this section, the essential properties of the three-dimensional four-fermion model (sometimes referred to as the Gross Neveu model (GNM_3)) will be reviewed. The Lagrangian of the model is

$$\mathcal{L} = \bar{\psi}_i(\partial\!\!\!/ + m_q)\psi_i - \frac{g^2}{2N_f}\left[(\bar{\psi}_i\psi_i)^2 - (\bar{\psi}_i\gamma_5\psi_i)^2\right], \tag{4.2}$$

where the index i runs over N_f fermion species. The problems in defining the γ matrices and Dirac spinors in $2+1$ dimensions are overcome by using the method given in [64]. ψ and $\bar{\psi}$ are taken to be four component spinors. The usual properties of the γ matrices still hold. m_q is the bare fermion mass.

To simplify both numerical and analytical work in this model scalar, σ, and pseudoscalar, π, auxiliary fields are introduced, so Eq. 4.2 becomes

$$\mathcal{L} = \bar{\psi}_i(\partial\!\!\!/ + m_q + \sigma)\psi_i - \frac{N_f}{2g^2}(\sigma^2 + \pi^2). \tag{4.3}$$

An identical generating function to that derived from the original Lagrangian can be recovered by a Gaussian integration over the auxiliary fields. At tree level the field σ is truly auxiliary, i.e. it has no dynamics. However, it acquires dynamical content from quantum effects arising from integrating out the fermions. Chiral symmetry breaking ($g > g_c^2$), in the chiral limit $m_q \to 0$, is now signalled by a non-vanishing expectation value, $\Sigma \equiv \langle\sigma\rangle$, for the scalar field. Σ serves as a convenient order parameter for the theory's critical point. From Eq. 4.3 it follows that the fermion gets a dynamically generated mass of $M \simeq \Sigma$.

This Lagrangian has a $U(1)$ chiral symmetry, although by setting the π fields to zero it becomes Z_2.

$$U(1) \quad : \quad \psi_i \to \exp(i\alpha\gamma_5)\psi_i \quad ; \quad \bar{\psi}_i \to \bar{\psi}_i \exp(i\alpha\gamma_5), \tag{4.4}$$

$$\mathrm{Z}_2 \quad : \quad \psi_i \to \gamma_5\psi_i \quad ; \quad \bar{\psi}_i \to -\bar{\psi}_i\gamma_5. \tag{4.5}$$

This symmetry is spontaneously broken whenever a non-vanishing condensate $\langle\bar{\psi}\psi\rangle$ is generated.

In three spacetime dimensions the following properties hold:

- For sufficiently strong coupling g^2 the model exhibits dynamical chiral symmetry breaking at zero temperature and density [64, 65].
- The spectrum of excitations contains both baryons and mesons, i.e. the elementary fermions and composite fermion–anti-fermion states.
- For $2 < d < 4$ the model has an interacting continuum limit [64, 65].
- When formulated on the lattice, the model has real Euclidean action even for chemical potential $\mu \neq 0$ [66], and hence can be simulated by standard Monte Carlo techniques.

This model is a useful toy model for understanding the behaviour of strongly interacting matter at high temperature and density since it displays much of the essential physics except for colour confinement.

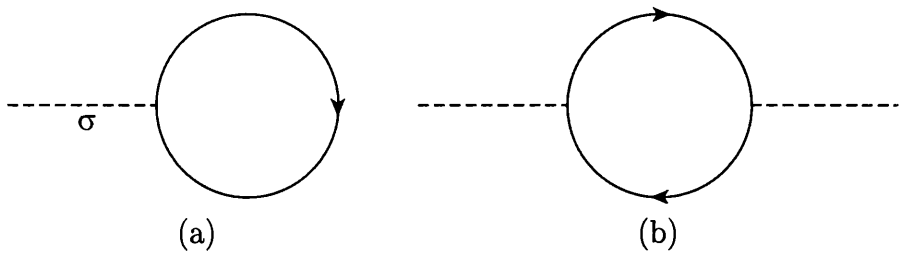

Figure 4.1: Leading order diagrams in Gross-Neveu model.

4.2 Mean field analysis at zero temperature and density

By taking the $N_f \to \infty$ limit, fluctuations around the saddle point solution are suppressed, this is equivalent to a mean field theory (MFT) treatment. Σ can be calculated using an expansion in inverse powers of N_f which associates a factor of N_f with each closed fermion loop and, in effect, $1/\sqrt{N_f}$ with each fermion scalar interaction vertex. To leading order, in the chiral limit $m_q \to 0$, only the tadpole diagram (see Fig. 4.1(a)) contributes to Σ, leading to the self-consistent Gap Equation [64]:

$$\frac{\Sigma}{g^2} = \int_p tr \frac{1}{i\not{p} + \Sigma}, \tag{4.6}$$

or, with a simple UV momentum-cutoff Λ we find a non-trivial solution $\Sigma \neq 0$, which breaks chiral symmetry if

$$\frac{1}{g^2} < \frac{1}{g_c^2} = \frac{2\Lambda}{\pi^2}. \tag{4.7}$$

Note that $\Sigma \to 0$ as $g^2 \to g_c^2$.

To the same leading order there is a correction to the scalar propagator (equal to g^2/N_f at tree level) from the bubble diagram (see Fig. 4.1(b)). Remarkably, the linear divergence in this diagram is cancelled by the divergence in the definition of g^2, leading to a closed-form expression which is finite when expressed in terms of Σ:

$$D_\sigma(k) = \frac{1}{N_f} \frac{2\pi\sqrt{k^2}}{(k^2 + 4\Sigma^2)\tan^{-1}\left(\frac{\sqrt{k^2}}{2\Sigma}\right)} \tag{4.8}$$

For the π field in the $U(1)$ model a similar expression is obtained, but with $(k^2+4\Sigma^2)$ in the denominator replaced by k^2. In the IR limit,

$$\lim_{k\to 0} D_\sigma(k) \propto \frac{1}{k^2 + 4\Sigma^2} \quad ; \quad \lim_{k\to 0} D_\pi(k) \propto \frac{1}{k^2}, \tag{4.9}$$

and hence σ resembles a fundamental boson with mass $m \equiv 2\Sigma$. Therefore the scalar is a weakly bound fermion–anti-fermion composite, whereas the π is a Goldstone mode. In the UV limit, we have

$$\lim_{k\to\infty} D_{\sigma,\pi}(k) \propto \frac{1}{\sqrt{k^2}}. \tag{4.10}$$

Thus the UV asymptotic behaviour is harder than that of a fundamental scalar $(1/k^2)$, but still softer than the $1/k^0$ corresponding to a non-propagating auxiliary field. This is down to the strong interaction between the fermion and the anti-fermion, since it causes diagrams corresponding to higher order corrections to be less divergent than expected by naive power counting.

The transition between chirally symmetric and broken phases at $g^2 = g_c^2$ defines a UV fixed point of the renormalisation group. It is characterised by non-Gaussian values for the critical exponents [65]:

$$\beta = \nu = \gamma = \eta = 1 \quad ; \quad \delta = 2 \tag{4.11}$$

The exponents obey certain consistency checks known as scaling and hyperscaling relations:

$$\gamma = \nu(2 - \eta) \quad ; \quad \beta = \frac{1}{2}\nu(d - 2 + \eta) \tag{4.12}$$

To derive these values in statistical physics [67], it is assumed that there is a single length scale, the correlation length, ξ, characterising all the important physics. With $\xi \sim \Lambda/\Sigma$, this is precisely the statement of renormalisability [65]. Corrections to these values are $\mathcal{O}(1/N_f)$ and calculable [65]. Indeed they are currently known to $\mathcal{O}(1/N_f^2)$ [68, 69, 70, 71], and when extrapolated to small values of N_f are supported by Monte Carlo estimates [72]. The continuum limit $g^2 \to g_c^2$ may be taken in either phase.

The deviation of the critical indices from the Gaussian model is related to the anomalous dimensions of the various composite fields in the model [65]. The universality class of the four-fermion model is not the standard, short range Ising model, but a Landau theory with long range forces [73, 74] and a specific value of the range parameter $R = d - 2$. These long ranges forces appear due to the existence of massless fermions that accompany the transition. However, they are not the only reason

for the differences seen in the four-fermion and Ising model critical exponents. The relationship between fermionic and scalar σ-models is more intricate and goes beyond naive universality arguments. Unlike scalar models with long range forces, where R is an external parameter, in the Gross-Neveu model R is generated dynamically by the fermions [73, 74]. Thus, the $N_f \to \infty$ limit of the Gross-Neveu model corresponds to a generalised Landau theory and the exponents of Eq. 4.11 replace the standard MFT ones. Another feature of these long range forces is that different N_f also produce different universality classes which are not related to the symmetry group.

4.3 Simulations

The semi-bosonized GNM_3 with $U(1)$ chiral symmetry used for the fermionic part of the lattice action is given by [75, 76]

$$\begin{aligned} S_{fer} &= \bar{\chi}_i(x) M_{ijxy} \chi_j(y) \\ &= \sum_{i=1}^{N_f} \left(\sum_{x,y} \bar{\chi}_i(x) \mathcal{M}_{xy} \chi_i(y) \right. \\ &\quad \left. + \frac{1}{8} \sum_x \bar{\chi}_i(x) \chi_i(x) \left[\sum_{\langle \tilde{x}, x \rangle} \sigma(\tilde{x}) + i\epsilon(x) \sum_{\langle \tilde{x}, x \rangle} \pi(\tilde{x}) \right] \right), \end{aligned} \tag{4.13}$$

where χ_i and $\bar{\chi}_i$ are Grassmann-valued staggered fermion fields defined on the lattice sites, the auxiliary fields σ and π are defined on the dual lattice sites, and the symbol $\langle \tilde{x}, x \rangle$ denotes the set of 8 dual lattice sites $\tilde{x}$ surrounding the direct lattice site x. The fermion kinetic operator $\mathcal{M}$ is given by

$$\mathcal{M}_{x,y} = \frac{1}{2} \sum_\nu \eta_\nu(x) \left[\delta_{y,x+\hat{\nu}} - \delta_{y,x-\hat{\nu}} \right] + m \delta_{x,y}, \tag{4.14}$$

where $\eta_\nu(x)$ are the Kawamoto-Smit phases $(-1)^{x_0+\cdots+x_{\nu-1}}$, and the symbol $\epsilon(x)$ denotes the alternating phase $(-1)^{x_0+x_1+x_2}$. The auxiliary fields σ and π are weighted in the path integral by an additional factor corresponding to

$$S_{aux} = \frac{N_f}{2g^2} \sum_{\tilde{x}} [\sigma^2(\tilde{x}) + \pi^2(\tilde{x})]. \tag{4.15}$$

The simulations were performed using a standard hybrid Monte Carlo algorithm without even/odd partioning, implying that simulation of N staggered fermions describes $N_f = 4N$ continuum species [75, 76]; the full symmetry of the lattice model in the continuum limit, however, is $U(N_f/2)_V \otimes U(N_f/2)_V \otimes U(1)$ rather than $U(N_f)_V \otimes U(1)$. At non-zero lattice spacing the symmetry group is smaller: $U(N_f/4)_V \otimes U(N_f/4)_V \otimes U(1)$. In the Z_2-symmetric model the π fields are switched off and M becomes real. In this case N staggered fermions describe $N_f = 2N$ continuum species. Further details of the algorithm and the optimisation of its performance can be found in [65, 75, 76].

Using point sources we calculated the zero momentum fermion (f) correlator at different values of the coupling $\beta \equiv 1/g^2$. In order to compare MEM to conventional spectroscopy we also estimated the fermion mass using a simple pole fit using the function

$$C_f(t) = A_f[e^{-M_f t} - (-1)^t e^{-M_f(L_t - t)}]. \tag{4.16}$$

Similarly, the zero momentum auxiliary π correlator was measured and its mass estimated using a cosh fit. The mesonic correlators are given by:

$$C_M(t) = \sum_{\mathbf{x},\mathbf{x_1},\mathbf{x_2}} \Phi(\mathbf{x}_1)\Phi(\mathbf{x}_2)W_M(\mathbf{x})G(\mathbf{x},t;\mathbf{x}_1,0)G^\dagger(\mathbf{x},t;\mathbf{x}_2,0), \tag{4.17}$$

where $W_M(\mathbf{x})$ is a staggered fermion phase factor which picks out a channel with particular symmetry properties i.e. $W_M(\mathbf{x}) = \epsilon(x)$ for the S channel and $W_M(\mathbf{x}) = 1$ for the PS channel. The function $\Phi(\mathbf{x})$ is either a point source $\delta_{\mathbf{x},(0,0)}$ or a staggered fermion wall source $\sum_{m,n=0}^{L_s/2-1} \delta_{\mathbf{x},(2m,2n)}$ [77]. In all the simulations we used point sinks. These correlators were fitted to a function $C_M(t)$ given by

$$C_M(t) = A[e^{-M_M t} + e^{-M_M(L_t-t)}] + \tilde{A}(-1)^t[e^{-\tilde{M}_M t} + e^{-\tilde{M}_M(L_t-t)}]. \tag{4.18}$$

Note that composite operators made from staggered fermion fields project onto more than one set of continuum quantum numbers. The first square bracket represents the "direct" signal with mass M_M and the second an "alternating" signal with mass $\tilde{M}_M$. Continuum quantum numbers for various mesonic channels are given in [78] – in this study we focus on the PS_{direct} channel, with $J^P = 0^-$. Although expected to be the tightest bound meson since it is the only one for which s-wave binding is available, as stressed in [75, 76, 78] this state does *not* project onto the Goldstone mode in the broken phase.

Symmetry	N_f	Volume	β	m_0	Configs
U(1)	4	32x32x48	0.55	0.005	22,600
U(1)	4	32x32x48	0.55	0.01	38,000
U(1)	4	32x32x48	0.55	0.02	43,000
U(1)	4	32x32x48	0.55	0.03	20,000
U(1)	4	32x32x48	0.55	0.045	19,300
U(1)	4	32x32x48	0.55	0.06	5,200
U(1)	4	32x32x48	0.65	0.01	60,000
U(1)	4	32x32x48	0.65	0.02	75,000
U(1)	36	24x24x32	0.55	0.01	6,500
U(1)	36	24x24x32	0.55	0.02	25,300
U(1)	36	24x24x32	0.55	0.03	10,900

Table 4.1: Broken phase data sets

4.3.1 Data sets

Tables 4.1–4.3 give the parameters for each of the data sets generated. They are split into three groups, firstly the broken phase pion, fermion and pseudoscalar (Table 4.1). Secondly the symmetric phase (Table 4.2), and finally the sigma in the broken phase (Table 4.3).

4.4 Theoretical spectral function form

4.4.1 Broken phase π, f and PS channels

In the broken phase, the π, f and PS are all expected to be simple poles. If J couples to a stable (i.e. zero width) bound state of mass M and strength A (i.e. $\langle 0|J|\overrightarrow{k}, M\rangle = A$), then $\rho(\omega) = (|A|^2/2M)\delta(\omega - M)$ (i.e. a delta function at M). Where $\rho(\omega)$ is now used as the spectral function previously defined as $f(\omega)$ in Chapter 2.

It is readily checked that the combination $\rho(\omega)/\omega^{d-2}$ is dimensionless for mesons in d-dimensions. This motivates the use of the default model $m_0(\omega) \propto \omega$ for GNM_3. This corresponds to the propagation of free massless fermions. For an asymptotically

Symmetry	N_f	Volume	β	Configs
Z_2	4	16x16x48	1.0	215,100
Z_2	4	24x24x24	1.0	121,500
Z_2	4	24x24x48	1.0	74,300
Z_2	4	32x32x48	0.92	41,200
Z_2	4	32x32x48	1.0	50,000
Z_2	4	32x32x48	1.25	39,000
Z_2	4	48x48x48	1.0	28,100
U(1)	4	32x32x32	1.0	31,000
U(1)	4	32x32x32	1.25	56,000
Z_2	36	24x24x32	1.125	22,200
Z_2	36	24x24x32	1.25	26,800

Table 4.2: Symmetric phase data sets

Symmetry	N_f	Volume	β	m_0	Configs
U(1)	4	32x32x24	0.65	0.01	1,741,600
U(1)	4	32x32x24	0.65	0.04	437,900
U(1)	4	24x24x24	0.70	0.01	1,000,000
U(1)	4	24x24x24	0.70	0.04	480,300
Z_2	4	24x24x24	0.65	0.00	1,100,000
Z_2	4	24x24x24	0.70	0.00	1,062,100

Table 4.3: Sigma data sets

free theory such as QCD, $\lim_{\omega\to\infty}\rho(\omega) = m_0(\omega)$ is expected (e.g. see Fig. 3.1). In GNM_3, however, this is not a constraint due to the UV behaviour being described by a renormalisation group fixed point with non-vanishing interaction strength [64, 65].

4.4.2 Symmetric phase

In the symmetric phase the momentum space propagator for the scalar channel in the large-N_f limit is

$$D_\sigma(k^2) \propto \frac{\mu^{d-2}}{(\sqrt{k^2})^{d-2} + \mu^{d-2}}, \tag{4.19}$$

where $2 < d < 4$ and μ is a dimensionful scale which increases as $(g_c^2 - g^2)^{\frac{1}{d-2}}$ (i.e. as an inverse correlation length). In three dimensions this implies

$$D_\sigma(t) \propto \mu \int_0^\infty dk \frac{\cos kt}{k+\mu} \equiv \mu \int_0^\infty d\omega \frac{\omega}{\omega^2+\mu^2} e^{-\omega t}, \tag{4.20}$$

hence the large-N_f prediction for the symmetric phase spectral function is

$$\rho_\sigma(\omega) \propto \frac{\mu\omega}{\omega^2+\mu^2}. \tag{4.21}$$

So in the asymptotic regime $lim_{\omega\to\infty}\rho_\sigma(\omega) \propto \omega^{-1}$ rather than the form of the default model $m_0(\omega) \propto \omega$.

At smaller scales we interpret ρ as describing a resonance whose central position and width are both $\mathcal{O}(\mu)$ and therefore increase as the coupling is reduced. In the IR limit $\lim_{\omega\to 0}\rho_\sigma(\omega)/\omega \propto$ constant.

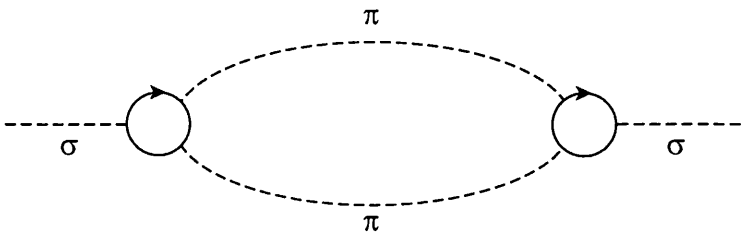

Figure 4.2: Contribution of 2π intermediate state to σ correlator.

4.4.3 Broken phase sigma

One of the major points of interest in using MEM on this model is the possibility of σ decay in the chirally broken phase. The σ is expected to be a weakly bound state of mass $M_\sigma \lesssim 2M_f$ (where M_f is the physical fermion mass), whereas, for the case of a continuous chiral symmetry, the pion mass m_π may be much smaller. If $2m_\pi < M_\sigma$ then the decay $\sigma \to 2\pi$ is allowed, and hence a feature should appear in the σ spectral function around the two pion threshold (defined below)

Fig. 4.2 shows diagrammatically the effects of the two pion intermediate state to the σ correlator. To leading order in $1/N_f$ the σ propagator taken from [65] is

$$D_\sigma(k^2) = \frac{1}{g^2}\frac{(4\pi)^{\frac{d}{2}}}{2\Gamma(2-\frac{d}{2})}\frac{M_f^{4-d}}{(k^2+4M_f^2)F(1,2-\frac{d}{2};\frac{3}{2};-\frac{k^2}{4M_f^2})}, \tag{4.22}$$

where for momenta $k \ll M_f$ the hypergeometric function, F, in the denominator may be approximated by $F \approx 1$. When the bare fermion mass $m > 0$, a similar expression for the pion propagator D_π is assumed, with $(k^2 + 4M_f^2)$ replaced by $(k^2 + m_\pi)$. The vertex $\Gamma_{\sigma\pi\pi}$ is assumed to arise from a single fermion loop, as shown in Fig. 4.2 If chiral symmetry is unbroken it is identically zero. It can be shown, using the bare vertex $-g/\sqrt{N_f}$, that

$$\Gamma_{\sigma\pi\pi} \simeq G_{\sigma\pi\pi} \frac{g^3 M_f^{d-3}}{\sqrt{N_f}}, \tag{4.23}$$

where $G_{\sigma\pi\pi}$ is a dimensionless d-dependent constant.

Now calculate D_σ in three dimensions including the effects of the two pion intermediate state.

$$D_\sigma^{-1}(k^2 \ll M_f^2) = \frac{g^2}{4\pi M_f}\left[k^2 + 4M_f^2 - \frac{G_{\sigma\pi\pi}^2}{N_f}\frac{M_f^3}{\sqrt{k^2}}\tan^{-1}\left(\frac{\sqrt{k^2}}{2m_\pi}\right)\right]. \tag{4.24}$$

So in addition to the pole at $k^2 = -4M_f^2$, there is now a contribution at $\mathcal{O}(1/N_f)$ to the timeslice correlation function given by

$$C_\sigma^{(1)}(t) \propto \frac{G_{\sigma\pi\pi}^2 M_f^3}{N_f}\int \frac{dk}{2\pi}\frac{e^{ikt}}{k(k^2+4M_f^2)^2}\tan^{-1}\left(\frac{k}{2m_\pi}\right), \tag{4.25}$$

The two pion threshold manifests itself via a branch cut in the inverse tangent from $k^2 = -4m_\pi^2$ to $\pm i\infty$. Taking the earlier approximate of $k^2 \ll M_f^2$ and integrating around the cut in the upper half plane, the following expression is obtained

$$C_\sigma^{(1)}(t) \propto \frac{G_{\sigma\pi\pi}^2}{32 N_f M_f}\int_{2m_\pi}^{\infty}\frac{d\omega}{\omega}e^{-\omega t} \tag{4.26}$$

from where it can be seen that

$$\rho_\sigma^{(1)} \propto \frac{G_{\sigma\pi\pi}^2}{32 N_f M_f}\frac{1}{\omega}\theta(\omega - 2m_\pi). \tag{4.27}$$

So the decay $\sigma \to 2\pi$ would produce a spectral feature at $\omega = 2m_\pi$ whose strength scales as $(N_f M_f m_\pi)^{-1}$. In principle this is testable by varying the simulation parameters N_f, g^2 and m, but on finite volumes, it will prove difficult to study the detailed form.

4.5 Results

The results will be split into the same three groups mentioned above. First the broken phase ($\beta < \beta_c \approx 1.0$) π, fermion and PS channels, all of which are expected to be simple poles/stable particles. Secondly the symmetric phase ($\beta > \beta_c$) where resonances with non-vanishing widths are expected. And finally the σ in the broken phase, where the two main issues to address are whether the σ is a bound state, and if it is possible to detect a signal for $\sigma \to \pi\pi$ decay.

4.5.1 Broken phase π, f and PS channels

A sample set of correlator data for the π, f and PS channels in the broken phase, i.e. $\beta < \beta_c \approx 1.0$, is shown on a log scale in Fig. 4.3. As expected all three appear to be a straight line, hence these channels are all dominated by a single particle pole.

Fig. 4.4 shows an example of the spectral functions obtained for the π, f and PS in the broken phase rescaled so they all fit the plot. All three particles appear as well-localised peaks, strongly suggesting, as expected, simple poles and hence are stable particle states. Table 4.4 gives a full comparison of the results obtained for the π, f and PS using standard one exponential fits and MEM which always agree within the errors stated.

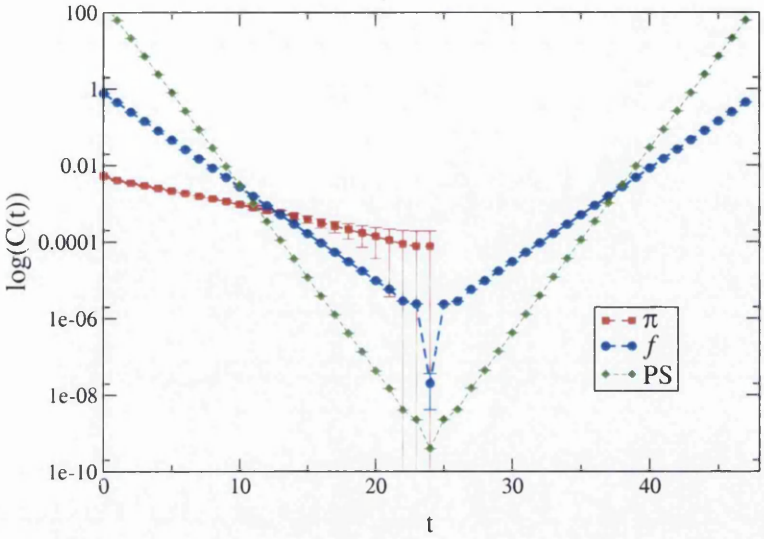

Figure 4.3: Propagators in three different channels from simulations of the U(1) model on a $32^2 \times 48$ lattice at $\beta = 0.55$, $m = 0.01$.

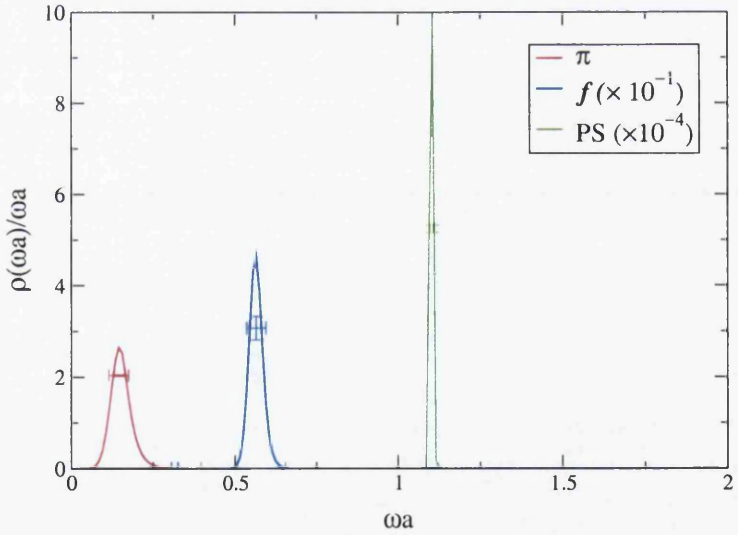

Figure 4.4: Three different channels from simulations of the U(1) model on a $32^2 \times 48$ lattice at $\beta = 0.55$ $m = 0.01$.

	N_f	Volume	β	m	Mass (1-exp)	Mass (MEM)	Area
π	4	$32^2 \times 48$	0.55	0.005	0.114(4)	0.112(6)	0.501(129)
	4	$32^2 \times 48$	0.55	0.01	0.168(5)	0.154(9)	0.176(15)
	4	$32^2 \times 48$	0.55	0.02	0.232(5)	0.231(7)	0.0617(98)
	4	$32^2 \times 48$	0.55	0.03	0.280(10)	0.263(15)	0.0351(37)
	4	$32^2 \times 48$	0.55	0.045	0.345(17)	0.324(14)	0.0188(11)
	4	$32^2 \times 48$	0.55	0.06	0.447(24)	0.424(8)	0.0101(20)
	4	$32^2 \times 48$	0.65	0.01	0.193(4)	0.187(8)	0.0810(78)
	4	$32^2 \times 48$	0.65	0.02	0.277(4)	0.267(6)	0.0289(19)
	36	$24^2 \times 32$	0.55	0.01	0.150(5)	0.144(18)	0.053(19)
	36	$24^2 \times 32$	0.55	0.02	0.238(6)	0.229(8)	0.0140(14)
	36	$24^2 \times 32$	0.55	0.03	0.287(10)	0.271(17)	0.0081(10)
f	4	$32^2 \times 48$	0.55	0.005	0.555(7)	0.556(4)	2.15(49)
	4	$32^2 \times 48$	0.55	0.01	0.564(1)	0.564(1)	2.37(3)
	4	$32^2 \times 48$	0.55	0.02	0.5853(7)	0.5858(13)	2.14(27)
	4	$32^2 \times 48$	0.55	0.03	0.599(1)	0.599(1)	2.06(5)
	4	$32^2 \times 48$	0.55	0.045	0.623(1)	0.623(1)	1.90(4)
	4	$32^2 \times 48$	0.55	0.06	0.644(2)	0.643(2)	1.63(8)
	4	$32^2 \times 48$	0.65	0.01	0.3978(8)	0.3965(13)	5.11(9)
	4	$32^2 \times 48$	0.65	0.02	0.4285(6)	0.4384(44)	4.10(33)
	36	$24^2 \times 32$	0.55	0.01	0.6796(3)	0.6796(3)	1.77(8)
	36	$24^2 \times 32$	0.55	0.02	0.6911(3)	0.6908(3)	1.72(7)
	36	$24^2 \times 32$	0.55	0.03	0.7025(4)	0.7023(5)	1.59(2)
PS	4	$32^2 \times 48$	0.55	0.005	1.0807(8)	1.0807(8)	164.3(6)
	4	$32^2 \times 48$	0.55	0.01	1.0973(8)	1.0979(7)	160(3)
	4	$32^2 \times 48$	0.55	0.02	1.1395(6)	1.1396(5)	147.2(5)
	4	$32^2 \times 48$	0.55	0.03	1.1715(11)	1.1716(11)	130(2)
	4	$32^2 \times 48$	0.55	0.045	1.2253(6)	1.2231(6)	119.1(9)
	4	$32^2 \times 48$	0.55	0.06	1.2693(13)	1.2691(2)	103(2)
	4	$32^2 \times 48$	0.65	0.01	0.7722(6)	0.7711(4)	426(32)
	4	$32^2 \times 48$	0.65	0.02	0.8362(5)	0.8381(45)	343(462)
	36	$24^2 \times 32$	0.55	0.01	1.3568(2)	1.3569(2)	50.1(3)
	36	$24^2 \times 32$	0.55	0.02	1.3806(2)	1.3808(2)	48.4(2)
	36	$24^2 \times 32$	0.55	0.03	1.4030(3)	1.4030(3)	45.5(3)

Table 4.4: Broken phase spectroscopy

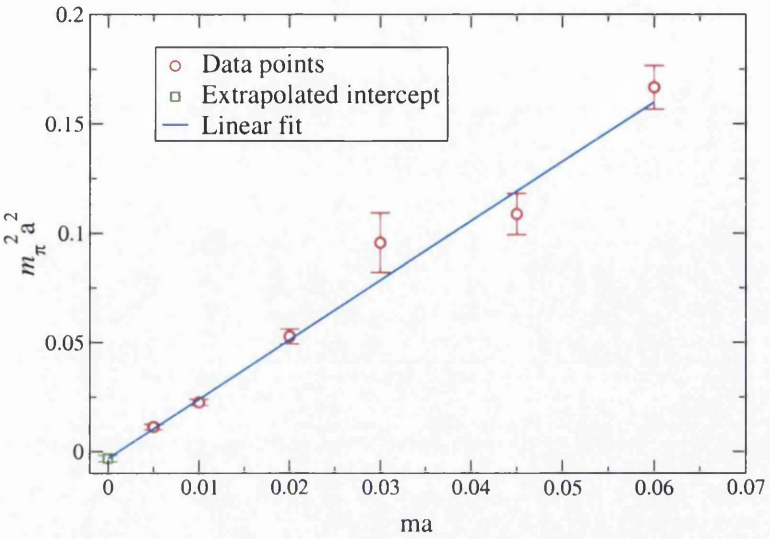

Figure 4.5: Pion mass m_π^2 vs. bare mass m for $\beta = 0.55$ showing evidence for the Goldstone nature of the π.

PCAC relation

Fig. 4.5 shows the scaling of the pion mass squared with the bare fermion mass. The line is a standard linear fit and the intercept is consistent with zero, hence the pion is consistent with the expected behaviour for broken chiral symmetry $m_\pi \propto \sqrt{m}$

PS binding energy

The PS is thought to be a weakly bound state consisting of two fermions. All the results tabulated in Table 4.4 are consistent with this (i.e. $M_{PS} < 2M_f$), and in fact due to the precision obtained on these numbers, it is possible to estimate the binding energy defined as $\Delta_M = 2M_f - M_{PS}$. The results for this calculation are tabulated in Table 4.5. For $N_f = 4$ $\Delta_M \approx 2.8\%$ of the bound state mass and for $N_f = 36$ $\Delta_M \approx 0.15\%$. This is consistent with the analytical expectation that $\Delta_M \propto 1/N_f$. Similar results in [78] observed that the PS wavefunction had considerably greater spatial extent for larger N_f, implying it is more weakly bound.

N_f	Volume	β	m	Δ_M (1-exp)	Δ_M (MEM)
4	$32^2 \times 48$	0.55	0.005	0.0293	0.0313
4	$32^2 \times 48$	0.55	0.01	0.0307	0.0301
4	$32^2 \times 48$	0.55	0.02	0.0311	0.0320
4	$32^2 \times 48$	0.55	0.03	0.0265	0.0264
4	$32^2 \times 48$	0.55	0.045	0.0207	0.0229
4	$32^2 \times 48$	0.55	0.06	0.0187	0.0169
4	$32^2 \times 48$	0.65	0.01	0.0234	0.0219
4	$32^2 \times 48$	0.65	0.02	0.0208	0.0387
36	$24^2 \times 32$	0.55	0.01	0.0024	0.0023
36	$24^2 \times 32$	0.55	0.02	0.0016	0.0008
36	$24^2 \times 32$	0.55	0.03	0.0020	0.0016

Table 4.5: Binding Energy in the PS channel

Changing the source/sink

In Fig. 4.6 the effect of using different meson sources following Eq. 4.17 using data from timeslices 1 – 8 is explored. The spectral functions have been rescaled so all fit on the same plot. When a wall is used at either sink or source, the signal is completely dominated by the bound state; however, for the point-to-point correlator there is a significant contribution out to $\omega a \approx 2.5$. Since we have discarded data from small timeslices we should not expect much quantitative information from the asymptotic form of $\rho(\omega)$ in this case; indeed, as $\omega \to \infty$ it decays much faster than either of the idealised forms $\rho_0(\omega)$ or $\rho_{UV}(\omega)$ discussed in § 4.4.1. Fig. 4.6 does however, provide a graphic illustration of the importance of choice of source in maximising the projection onto the ground state.

4.5.2 Symmetric phase

In the symmetric phase, $\beta > \beta_c$, both the Z_2 (σ channel only) and $U(1)$ (σ and π channels) models have been considered. It proved considerably easier to simulate

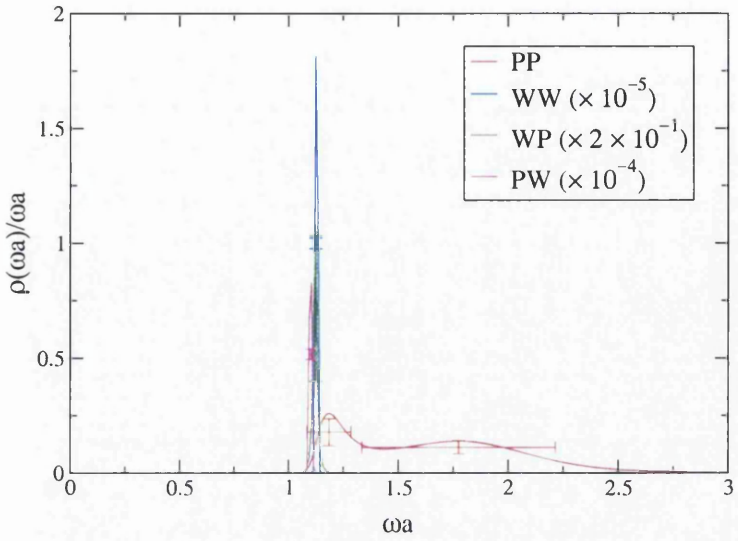

Figure 4.6: PS channel on $32^2 \times 48$ lattice at $\beta = 0.55$ $m = 0.01$ using correlators with different combinations of wall and point source and sinks

this phase for the Z_2 model, the $U(1)$ requiring a much smaller molecular dynamics timestep. The Z_2 σ correlator data in Fig. 4.7 shows clearly that a single-exponential fit will not work in this channel. The correlators become almost flat at large t

The expected form for the symmetric phase resonance is given in Eq. 4.21. In order to identify spectral features which are not simple poles, it is important to understand the systematic effects. Fig. 4.8 shows the effects of altering the timeslices used in the fit. Data for $t \geq 12$ has been discarded to avoid finite volume effects (actually non-zero temperature). In all cases there is a broad spectral feature at $\omega a \approx 0.5$ whose width increases as smaller timeslices are included. If the divergence as $\omega \to 0$ is taken to be an artifact and therefore ignored (similar features at $\omega \to 0$ have been seen in other MEM studies [79]), then the shape appears qualitatively similar to the large N_f prediction in § 4.4.2. The fact that the shape of this resonance in the massless phase is sensitive to the data at small times is slightly counter-intuitive, but

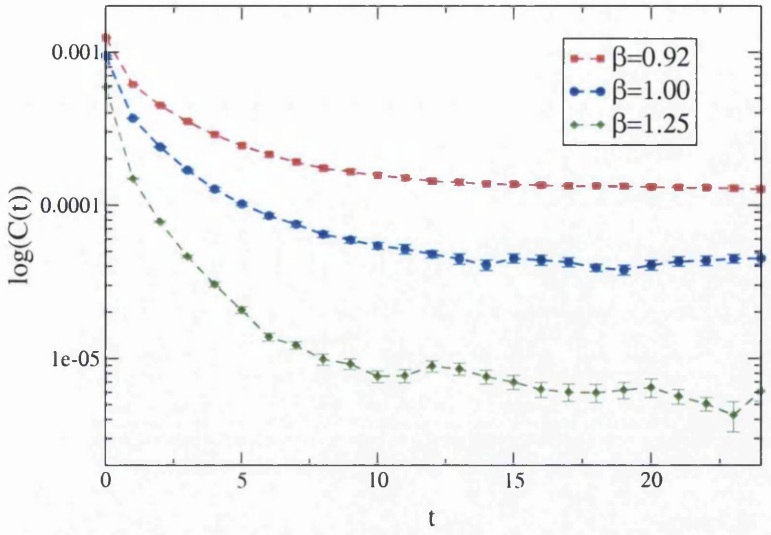

Figure 4.7: σ correlator for 3 different couplings in the chirally symmetric phase on a $32^2 \times 48$ lattice.

is consistent with the observation in [78] that the extraction of the physical scale, i.e. the resonance width μ, from timeslice correlator data actually depends on corrections to the expected power-law falloff $(\mu t)^{-2}$ at small values of μt. Notice also that as predicted $\lim_{\omega\to\infty} \rho(\omega)/\omega \to 0$ in contrast to an asymptotically free theory such as QCD.

The next systematic test was to change the number of configurations used in the fit to show up whether this width is purely due to insufficient statistics. Fig. 4.9 shows the effect on the peak of changing the number of configurations from $\mathcal{O}(10000)$ up to $\mathcal{O}(40000)$. The width of this feature remains fairly stable, although both the central position and area under the peak vary slightly, but this only supports the view that this is not a simple pole (see Table 4.6).

Fig. 4.10 shows the spectral functions obtained from three different couplings. The normalisation of the results is distorted by the artifact at $\omega \to 0$, the curves have been

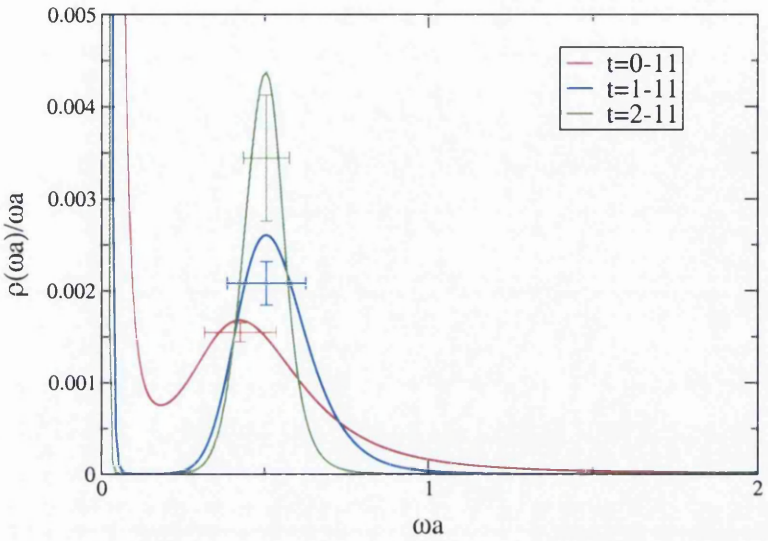

Figure 4.8: Bryan image of $\rho(\omega)/\omega$ vs. ω in the σ channel at $\beta = 1.25$ on a $32^2 \times 48$ lattice, showing 3 different time windows.

N_{cfg}	Central Position	Area	Width
19500	0.325(734)	0.032(32)	0.308(691)
29250	0.375(62)	0.062(40)	0.311(17)
39000	0.415(69)	0.091(51)	0.226(663)

Table 4.6: Analysis of resonance with changing N_{cfg} in the σ channel at $\beta = 1.25$. The smallest N_{cfg} seen in Fig. 4.9 does not appear in this table because there is no peak to analyse in this case.

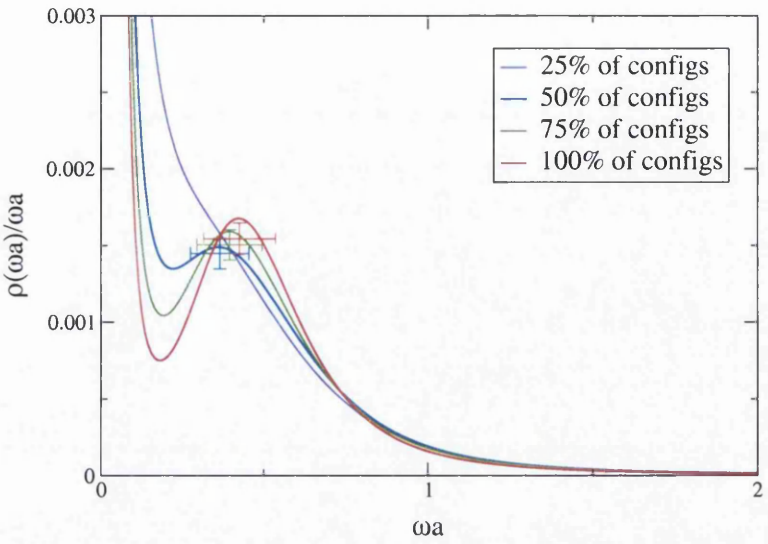

Figure 4.9: The same as Fig. 4.8 using fits from timeslices 0 – 11, showing the effects of varying the amount of data.

rescaled so that the height of each peak is the same. As β increases (i.e. coupling decreases) both the central position and width of the peak becomes larger. This is consistent with the predictions in § 4.4.2, which stated that both are proportional to a single scale, μ, which increases with β.

Finally the symmetric phase U(1) model is considered, where there are both σ and π channels. The two should be physically indistinguishable for $\beta > \beta_c$ and indeed for large ω at least this is the case in Fig. 4.11. However, for small ω there is a large disparity between the two β values, $\rho(\omega)$ appears to diverge for $\beta = 1.0$, but tend smoothly to 0 for $\beta = 1.25$. As can be seen from Fig. 4.12 the large t behaviour of these correlators is not under proper control with the statistical sample obtained. Another point to note is that in both cases there is more power in the σ channel at small ω. The only real conclusions which can be drawn from this analysis is that a full understanding of the systematics in this regime is still lacking. However, in

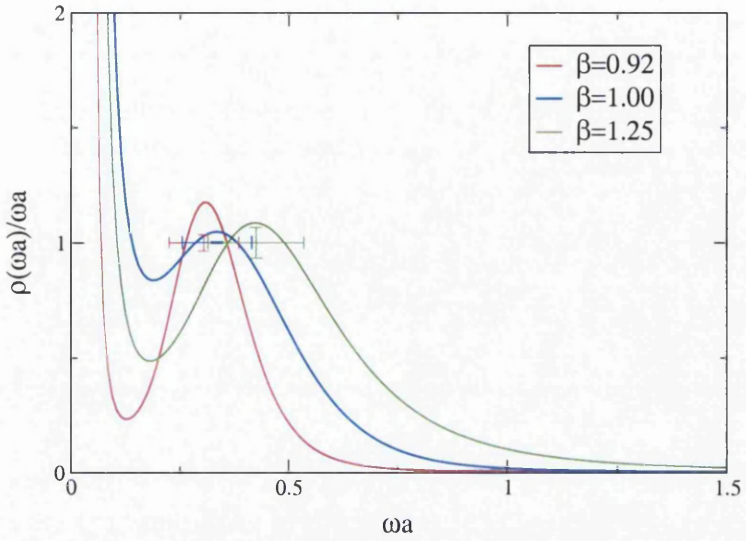

Figure 4.10: Rescaled Bryan image of $\rho(\omega)/\omega$ in the symmetric phase from timeslices 0 – 11, for three different couplings.

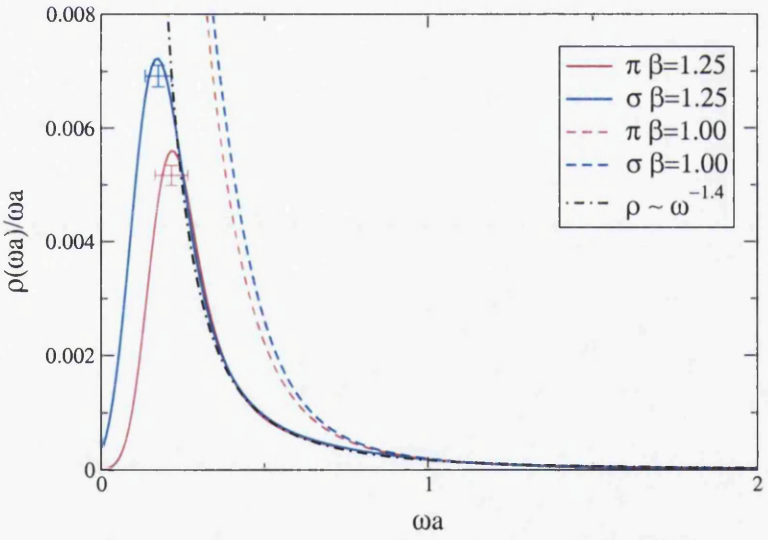

Figure 4.11: Bryan image of $\rho(\omega)/\omega$ in both σ and π channels from simulations of the U(1) model on a 32^3 lattice. The dashed-dotted line is of the form $\omega^{-1.4}$ close to the large-N_f prediction of ω^{-1}.

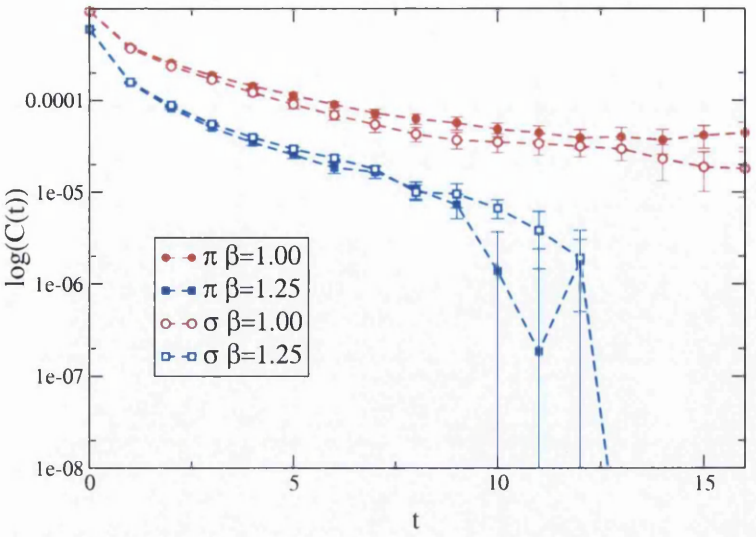

Figure 4.12: σ and π timeslice correlators from simulations of the U(1) model on a 32^3 lattice.

the large-ω regime, the behaviour of the spectrum is close to that of the large N_f prediction of $\rho(\omega) \propto \omega^{-1}$. The dash-dotted line included in Fig. 4.11 is of the form $\rho(\omega) \propto \omega^{-1.4}$.

To summarise:

- There is encouraging evidence that MEM is capable of identifying a resonance with a non-zero width.
- The properties are semi-qualitatively consistent with the theoretical expectations
- Uncertainties remain about the $\omega \to 0$ regime which would probably require significantly larger lattices in the time direction.

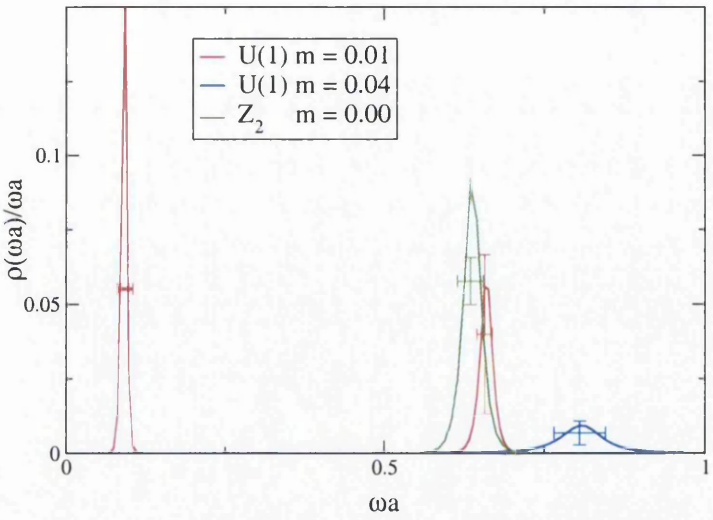

Figure 4.13: Rescaled Bryan image of $\rho(\omega)/\omega$ in the σ channel from timeslices 0 – 11 at $\beta = 0.65$, for two different masses in the U(1) model on a $32^2 \times 24$ lattice, and for $m = 0$ in the Z_2 model on a 24^3 lattice.

4.5.3 Broken phase sigma

Disconnected fermion lines are automatically included in the σ correlator since it is modelled by an auxiliary field. The main physical issues to address here are whether the σ is a bound state, and if there is a signal for $\sigma \to \pi\pi$ decay. Fig. 4.13 shows the spectral functions from 3 simulations, the $U(1)$ model with 2 different values of the bare fermion mass, m_q and the Z_2 model ($m_q = 0$). Very large statistical samples were used for this study, see Table 4.3.

Vacuum subtraction

Since σ shares the same quantum numbers as the vacuum, before analysing the correlator data it is necessary to subtract the vacuum term,

$$C_{\text{vac}} = \sum_{\vec{x},t} \langle \sigma(\vec{0},0) \rangle \langle \sigma(\vec{x},t) \rangle. \tag{4.28}$$

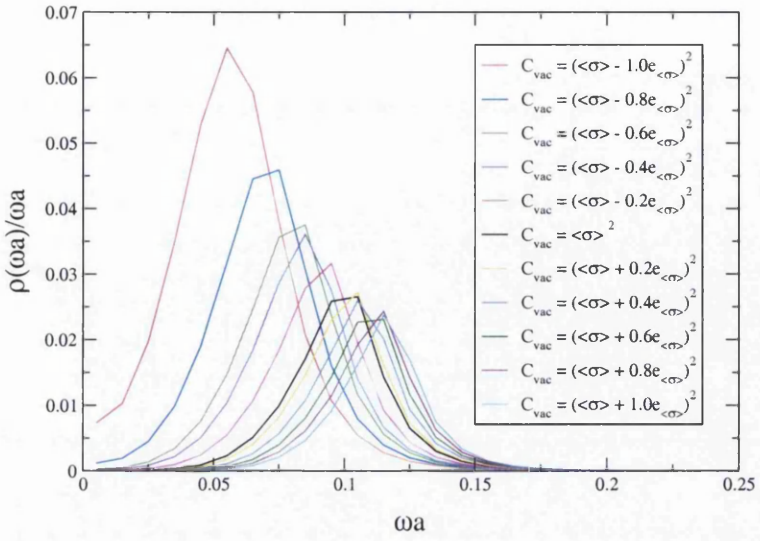

Figure 4.14: Variation of artifact with vacuum subtraction constant C_{vac}

Due to the statistical fluctuations this is a hard procedure to implement. In this section, the variation of the sharp spike centered at around $\omega a = 0.1$, which we believe is due to our uncertainty in the vacuum subtraction, is tested as we alter the value of the vacuum by one standard deviation ($e_{\langle\sigma\rangle}$) above and below. Fig. 4.14 and Table 4.5.3 show the analysis of this sharp spike graphically and numerically respectively. As the value of the vacuum subtracted is varied the artifact varies significantly while the σ peak remains consistent within the errors.

Width analysis

Similar to the procedure used for the symmetric phase resonances, the width of the σ resonance has been analysed at varying numbers of configurations. This is to eliminate the possibility that the large width seen on the spectrum is statistical. The results shown graphically in Fig. 4.15 and numerically in Table 4.8 shows the width remains stable within errors even when the statistical sample is dramatically increased

C_{vac}	Artifact			Sigma		
	Mass	Area	Width	Mass	Area	Width
$-1.0e_{\langle\sigma\rangle}$	0.056(113)	0.003(16)	0.020(49)	0.657(9)	0.0014(1)	0.051(27)
$-0.8e_{\langle\sigma\rangle}$	0.069(64)	0.0021(47)	0.018(17)	0.658(7)	0.00142(7)	0.049(19)
$-0.6e_{\langle\sigma\rangle}$	0.079(54)	0.0017(27)	0.018(15)	0.658(6)	0.00142(5)	0.048(16)
$-0.4e_{\langle\sigma\rangle}$	0.083(51)	0.0015(22)	0.018(14)	0.659(6)	0.00142(5)	0.047(16)
$-0.2e_{\langle\sigma\rangle}$	0.090(46)	0.0013(17)	0.017(12)	0.659(6)	0.00141(5)	0.046(15)
0	0.098(42)	0.0011(12)	0.016(10)	0.660(6)	0.00141(6)	0.045(13)
$+0.2e_{\langle\sigma\rangle}$	0.101(41)	0.0010(11)	0.016(10)	0.660(6)	0.00140(4)	0.045(13)
$+0.4e_{\langle\sigma\rangle}$	0.105(39)	0.0010(9)	0.016(9)	0.660(5)	0.00140(4)	0.045(12)
$+0.6e_{\langle\sigma\rangle}$	0.109(37)	0.0009(8)	0.016(8)	0.661(5)	0.00140(4)	0.044(12)
$+0.8e_{\langle\sigma\rangle}$	0.112(36)	0.0009(8)	0.015(8)	0.661(5)	0.00140(4)	0.044(12)
$+1.0e_{\langle\sigma\rangle}$	0.115(35)	0.0008(7)	0.015(7)	0.661(5)	0.00139(4)	0.043(12)

Table 4.7: Detailed analysis of the artifact and physical sigma peak as the C_{vac} is varied up to one standard deviation either side. See Figs. 4.13 and 4.14.

N_{cfg}	Central Position	Area	Width
435,400	0.642(12)	0.00169(24)	0.058(26)
870,800	0.659(6)	0.00155(4)	0.039(9)
1,306,200	0.635(2)	0.00173(3)	0.048(16)
1,741,600	0.660(6)	0.00141(4)	0.045(13)

Table 4.8: Analysis of σ resonance for U(1) $\beta = 0.65$ $m_q = 0.01$ on $32^2 \times 24$ lattice with changing N_{cfg}.

suggesting that the width of the σ is not purely statistical, i.e. the σ is a resonance and not a simple pole.

In contrast to the large-N_f predictions in § 4.4.3 that $\rho(\omega)$ should be sharply cut off on the low-ω side, but fall away more slowly on the high-ω side due to an $f\bar{f}$-continuum, the shape of the resonance is roughly symmetric, unlike those for the PS in Fig. 4.6.

The central value for the peak in the $U(1)$ $m_q = 0.01$ data set is lower than the corresponding PS state (see Table 4.4), which is at 0.77. The $f\bar{f}$ threshold in the case is at 0.793(3), which is well above the point where $\rho(\omega)/\omega$ appears to fall to zero. It

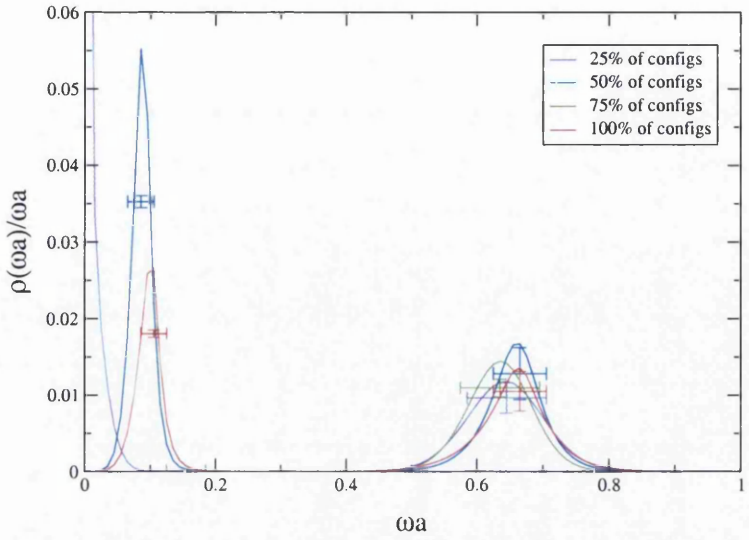

Figure 4.15: U(1) $32^2 \times 24$ $\beta = 0.65$ $m = 0.01$ σ spectral functions for different numbers of configurations. The peak which sometimes appears around $\omega a = 0.1$ is an artifact due to difficulties with vacuum subtraction (see § 4.5.3)

is therefore deduced that for finite N_f, that the σ is a more tightly bound state than the PS meson for which there are no disconnected fermion line contributions.

Sigma decay

Unfortunately, there is no sign of any spectral feature around the two pion threshold, which would be expected at $\omega a \simeq 0.38$ for $m_q = 0.01$ and $\omega a \simeq 0.75$ for $m_q = 0.04$ (see § 4.4.3 for derivation of the two pion threshold). Recall that the condition for $\sigma \to \pi\pi$ is $M_\sigma < 2m_\pi$ so in the former case it is certainly possible. A recent study of the O(4) sigma model by Ishizuka and Yamazaki has claimed to see evidence of $\sigma \to \pi\pi$ decay [80].

4.6 Conclusions

Studies of theories beyond quenched QCD at zero temperature will require greater sophistication than the current single- and multi-exponential fits, which assume a spectral density function consisting of a series of isolated poles. It is clear from the study of the symmetric phase and σ in the broken phase that these techniques will provide poor results. This is the first attempt at using the Maximum Entropy Method to a lattice model with dynamical fermions.

In the chirally broken phase of the model, the elementary fermion f, the simplest $f\bar{f}$ bound state and the Goldstone boson π have all been shown to be sharp spectral features (i.e. simple poles). This confirms the findings of earlier studies [75, 76, 78]. Estimates for the meson binding energy have also been made for the first time.

In the chirally symmetric phase, a broad resonance, whose features agree qualitatively with the large-N_f predictions, has been identified.

The first quantitative study of the σ channel in the chirally broken phase has also been made and found that it is more tightly bound than the conventional PS meson due to the additional contribution from disconnected fermion line diagrams. Unfortunately there was no evidence of any feature at the two pion threshold, and therefore no evidence for $\sigma \rightarrow \pi\pi$ decay.

The philosophy of MEM is to make maximum possible use of the data available, unlike single-exponential fitting where the time window is chosen to coincide with a plateau in the effective mass plot. The main problem faced in the studies above has been associated with the upper end of the time window. In an attempt to avoid any

finite volume effects associated with including the large timeslices in the fit, control over the $\omega \to 0$ limit has been sacrificed. This can easily be seen in many of the plots presented where the spectral density function appears to diverge at $\omega = 0$. There were also artifacts in some cases of the broken phase σ which arose due to difficulties with the vacuum subtraction.

Chapter 5

Dynamical QCD

The results for the spectral functions obtained from the UKQCD Collaboration's dynamical fermion simulations are presented in this chapter. The quenched simulation has been included here to provide a direct comparison with the dynamical results. We begin by detailing the simulation parameters.

5.1 Simulation parameters

In addition to the three dynamical data sets and the corresponding quenched simulation (where the sea quark mass is infinite) forming a matched ensemble, a further dynamical data set at lighter sea quark masses was simulated. The greatest effects due to the inclusion of the dynamical fermions are seen when the sea quark mass is made as light as possible, ideally in the vicinity of the up and down quark masses. However, since the computational time required for simulations increases as the sea quark mass decreases it is currently not feasible with current computer resources. The lightest sea quark mass chosen here represents the smallest quark mass at which meaningful statistics could be achieved within an acceptable period of time.

β	c_{sw}	κ_{sea}	κ_{val}	No. config.
5.29	1.9192	0.1340	0.1335, 0.1340, 0.1345, 0.1350	101
5.26	1.9497	0.1345	0.1335, 0.1340, 0.1345, 0.1350	102
5.2	2.0171	0.1350	0.1335, 0.1340, 0.1345, 0.1350	151
5.93	1.82	Quenched	0.1327, 0.1332, 0.1334, 0.1337, 0.1339	623
5.2	2.0171	0.1358	0.1340, 0.1345, 0.1350, 0.1355	101

Table 5.1: Simulation parameters for all the data sets, the last simulation is unmatched

Gauge configurations were generated with two degenerate flavours of $\mathcal{O}(a)$ improved dynamical Wilson fermions using the Hybrid Monte Carlo algorithm [81] on the Cray T3E supercomputer in Edinburgh. The implementation and verification of the code was described in [82, 83] and a summary of the algorithm details was reported in [84]. The dynamical gauge configurations were separated by 40 trajectories (a figure decided on after a study of the autocorrelation times measured for the plaquette on every trajectory [84, 82, 85]). The matched quenched gauge configurations were generated by the hybrid over-relaxed algorithm with the compound sweep ratio of 7:1, over-relaxed to Cabbibo-Marinari sweeps [86]. The separation for the gauge configurations used for measurements was 700 compound sweeps.

Quark propagators were generated using $\mathcal{O}(a)$ improved Wilson fermions. Correlators were constructed from fuzzed propagators for degenerate combinations of κ_{val}. For the quenched simulation, degenerate and non-degenerate meson correlator combinations were generated for three values of the hopping parameter. A further two κ_{val} values were added to the simulation to achieve a lower m_{PS}/m_V mass ratio, which was more comparable to the lighter dynamical simulations.

The simulation parameters for all the data sets is given in Table 5.1, all of which

were carried out on a $16^3 \times 32$ lattice. A lattice of this size was deemed necessary, following the finite volume investigation in [84], to keep the finite size effects to a minimum as the sea quark mass (and hence lattice spacing, a) was reduced. In order to ensure a large enough spatial volume to accommodate baryons ($\gtrsim$ 1.5fm) at this lattice size, a coarse lattice spacing of $a \gtrsim 0.09$fm is required. Therefore simulations must be performed at low β values. The dynamical β value was determined to be as low as possible while remaining within the parameter range where a valid non-perturbative estimate of the clover coefficient had been determined.

The fully non-perturbatively $\mathcal{O}(a)$ improved values for c_{sw} were used for all the dynamical simulations. This was determined by ALPHA Collaboration in [87] and is given by

$$c_{sw}^{\text{dynam}} = \frac{1 - 0.454g_0^2 - 0.175g_0^4 + 0.012g_0^6 + 0.045g_0^8}{1 - 0.720g_0^2}. \tag{5.1}$$

This is valid for β values as low as 5.2, the minimum value included in the simulations. Hence residual lattice artifacts are expected to be of $\mathcal{O}(a)$, which on the coarse lattices used in these simulations could still be significant. For the quenched simulation, the clover coefficient was determined by the SCRI Collaboration [88]

$$c_{sw}^{\text{quen}} = \frac{1 - 0.6084g_0^2 - 0.2015g_0^4 + 0.03075g_0^6}{1 - 0.8743g_0^2}, \quad \beta \geq 5.7. \tag{5.2}$$

This result extends the analysis of the ALPHA Collaboration to lower values of β.

The values of r_0 have been obtained by the UKQCD collaboration [11] and are

$(\beta, \kappa_{\text{sea}})$	r_0/a	$a[fm]$
(5.20, .1350)	$4.754(40)(^{+2}_{-90})$	$0.1031(09)(^{+20}_{-1})$
(5.26, .1345)	$4.708(52)(^{+45}_{-50})$	$0.1041(12)(^{+11}_{-10})$
(5.29, .1340)	$4.813(45)(^{+35}_{-84})$	$0.1018(10)(^{+20}_{-7})$
(5.93, 0)	$4.714(13)(^{+0}_{-18})$	$0.1040(03)(^{+4}_{-0})$

Table 5.2: Summary of r_0 and a for the matched data set, the errors quoted are statistical and systematic respectively.

summarised, along with the lattice spacing derived from r_0 to illustrate the level of matching achieved in Table 5.2

5.2 MEM technicalities

We now turn to the details on the application of the MEM technique to this data. The channels to be analysed are the axial temporal and spatial, scalar, pseudoscalar and vector. The default model used is of the form

$$m(\omega) = m_0\omega^2 \tag{5.3}$$

motivated by the asymptotic behaviour of the spectral function at large ω. The frequency space was discretised into $n_\omega = 600$ parts with a separation of $\omega a = 0.01$. The integration over α was done for all probabilities $> 1\%$ of the maximum

5.3 Results

The analysis of the data sets will be split into four sections. Firstly the quenched data will be analysed, followed by the matched ensemble including the quenched simulation for comparison. Then the lightest sea quark mass data is studied. An additional discussion on the non-singlet scalar meson will then be given.

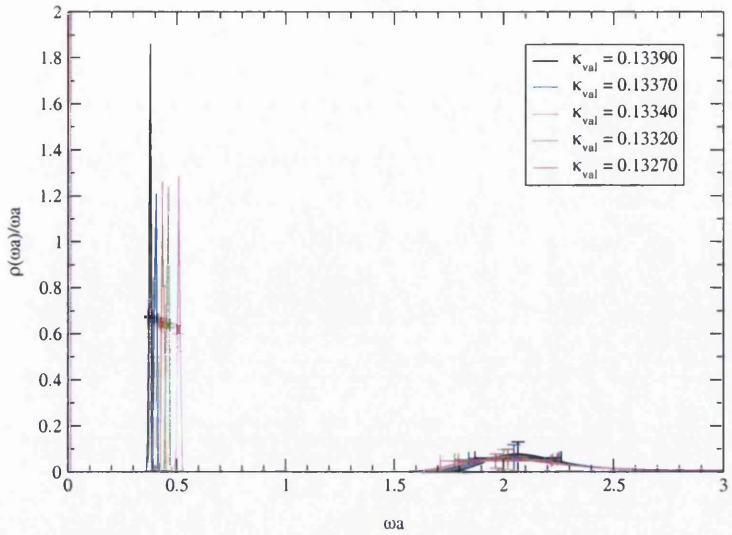

Figure 5.1: UKQCD $\beta = 5.93$ (Quenched) Axial Temporal channel t=0-14

5.3.1 Quenched data set

The first data analysed and presented here is the quenched data set for all five valence quark masses in degenerate pairs. This will give a base for comparison when the dynamical data is analysed. The effects of changing the quark mass can be seen in these graphs and should give us some insight into whether any changes seen in the matched dynamical data set (see § 5.3.2) are from the change of sea quark mass, or simply equivalent to a change in the valence quark mass in the quenched approximation.

The spectral functions for the axial temporal, pseudoscalar and vector channels all display a common structure: a sharp peak at low energy and a broad bump at higher energy (two bumps in the vector case). In the quenched approximation the ground state peaks in these channels should have zero width, but due to the finite number of configurations (see the tests performed on the delta function test data in § 3.2.5) and possibly the discrete nature of the lattice in the temporal direction these unphysical widths are produced. The integrated strength of the peak, however, can

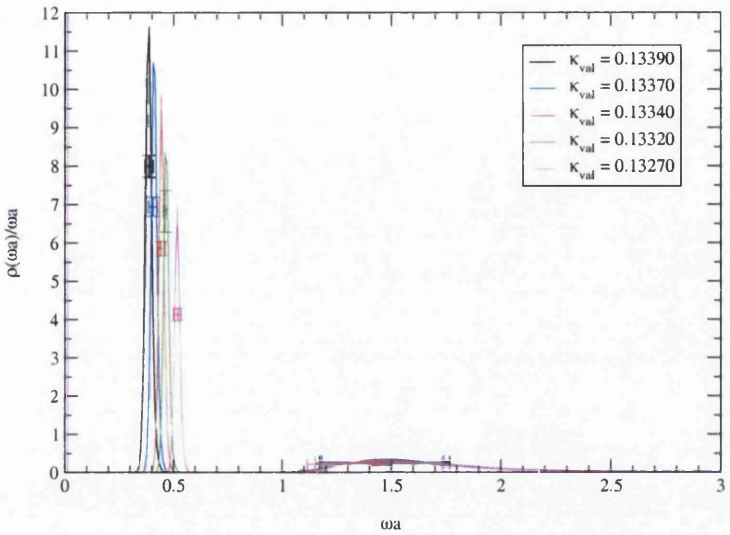

Figure 5.2: UKQCD $\beta = 5.93$ (Quenched) Pseudoscalar channel t=0-14

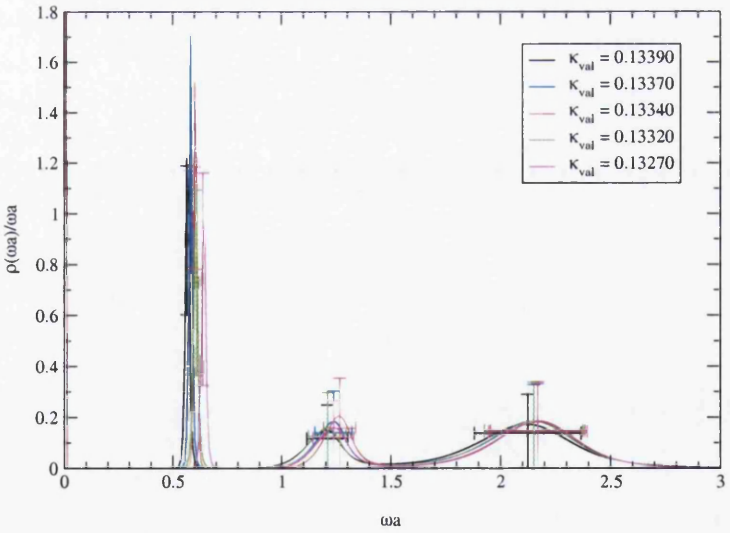

Figure 5.3: UKQCD $\beta = 5.93$ (Quenched) Vector channel t=0-14

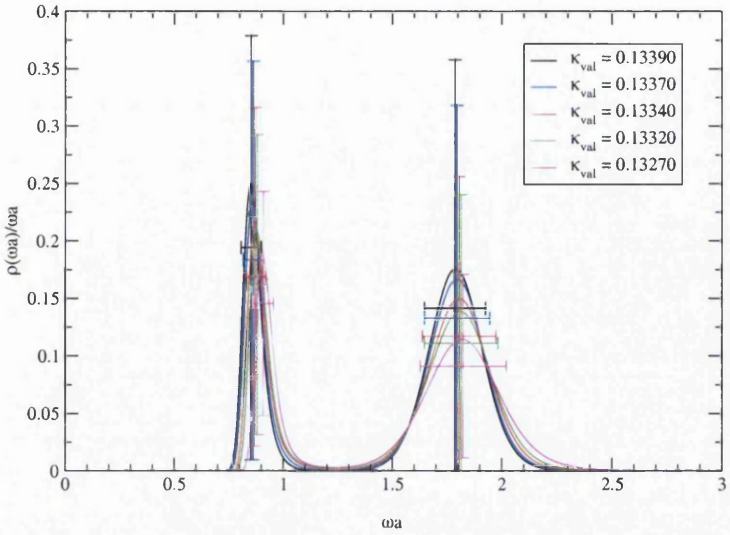

Figure 5.4: UKQCD $\beta = 5.93$ (Quenched) Axial spatial channel t=0-14

be related to the physical decay constant of the mesons. From the error bar analysis on the broad peaks, very little statistical significance should be made of their shape since the error bars are a similar size to the height of the peak.

The effects of changing the quark mass are as expected, the mass of the meson (central position of the peak) increases with the valence quark mass (i.e. as κ_{val} decreases)

The remaining two channels, the axial spatial and scalar, are difficult to fit using traditional exponential fits. In the spectral functions of these channels (Figs. 5.4 and 5.5) there is perhaps some evidence to suggest why. The two peaks found are both broad, even the ground state (compare with the sharp ground state peak found in the other channels), which makes the assumption of a delta function form for these inappropriate. The error bars are large though, especially in the axial spatial channel. The scalar channel is discussed in more detail in § 5.4

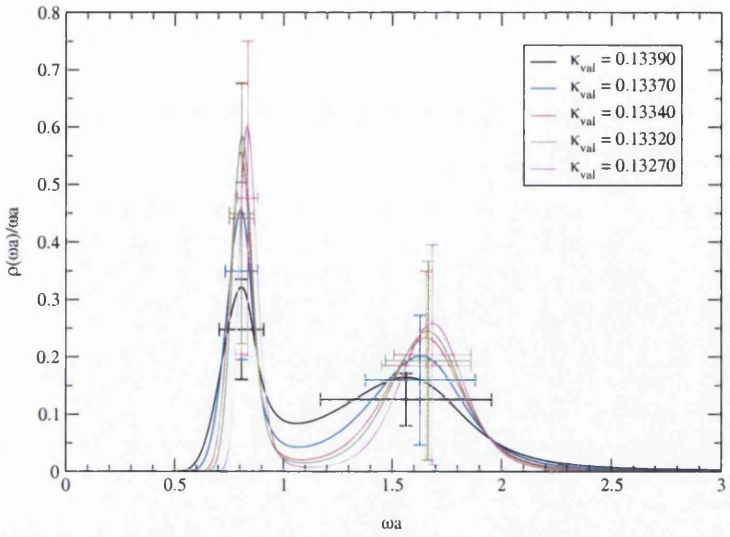

Figure 5.5: UKQCD $\beta = 5.93$ (Quenched) Scalar channel t=0-14

β	κ_{sea}	κ_{val}	am_{PS} (2-exp)	am_{A_0} (MEM)
5.20	0.13500	0.13500	$0.405\ ^{+4}_{-5}$	$0.414\ ^{+9}_{-9}$
5.26	0.13450	0.13450	$0.509\ ^{+2}_{-2}$	$0.498\ ^{+6}_{-6}$
5.29	0.13400	0.13400	$0.577\ ^{+2}_{-2}$	$0.590\ ^{+8}_{-8}$
5.93	0.0	0.13390	$0.356\ ^{+1}_{-1}$	$0.378\ ^{+9}_{-9}$

Table 5.3: Comparison with UKQCD axial temporal masses

5.3.2 Matched Unitary Set

Next the results from the matched (fixed lattice spacing a) unitary ($\kappa_{\text{sea}} = \kappa_{\text{val}}$) data set are reviewed. This data was chosen because it isolates the effects of changing the quark mass, so any change in the spectral functions is presumably due to the quark mass rather than any $\mathcal{O}(a^2)$ lattice artefact. These dynamical quark effects should become more prominent as smaller, more physical quark masses are approached, i.e. as κ_{sea} increase. One matched quenched spectrum has been included on each of these graphs for comparison.

Fig. 5.6 presents the spectral functions for the axial temporal channel. Since the

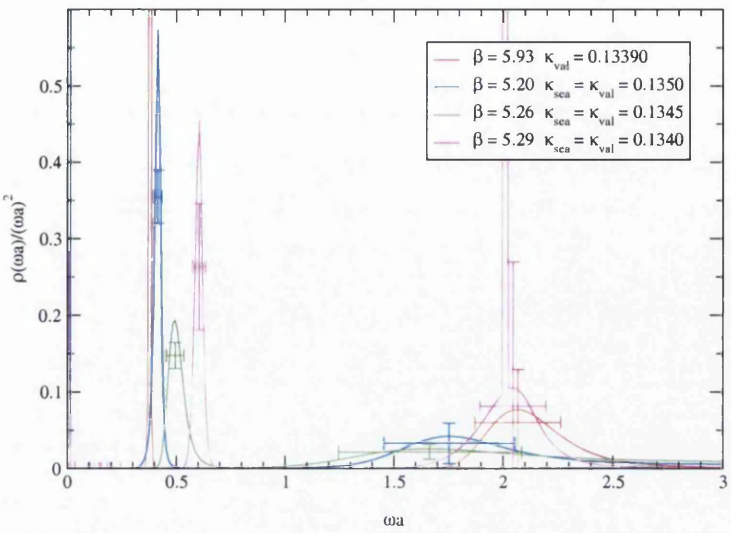

Figure 5.6: UKQCD Axial Temporal t=0-14 matched unitary data sets

axial temporal ground state shares the same quantum numbers as the pseudoscalar the comparison in Table 5.3 is between the two-exponential PS fits and MEM's axial temporal analysis. Details of the UKQCD two-exponential fits are discussed in [11]. MEM results from local source and sink is in agreement within the 2σ level with the two-exponential results which have used combinations of local and fuzzed source and sink to improve the signal. Note that the errors quoted in Table 5.3 and subsequent comparisons are statistical only, the systematic errors are assumed to be of the same order of magnitude again As lighter quark masses are approached the ground state peak shifts to lower energies. The mass is therefore decreasing just as can be seen in Fig. 5.1 for the quenched data as the value of κ is increased. The resonance centered around $\omega a \approx 2.0$ varies only slightly and it is very difficult to make any conclusions given the errors on this feature.

The axial spatial results are presented in Fig. 5.7. This is an extremely difficult channel to extract any information from, so there are no two-exponential values to

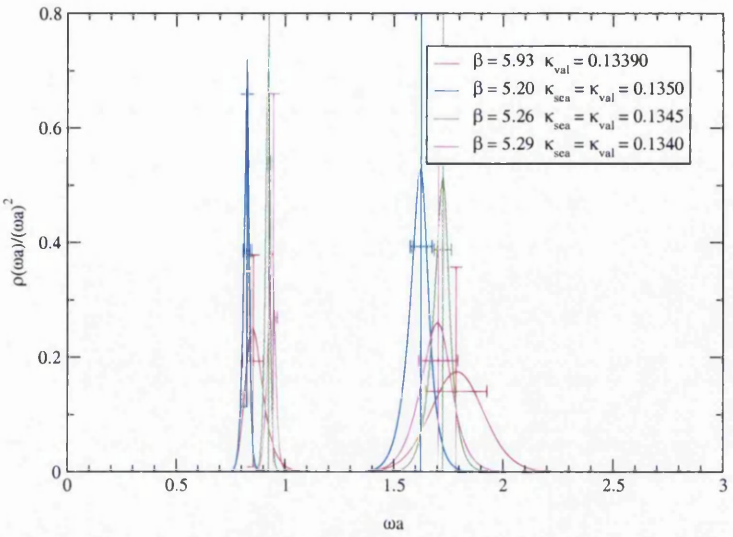

Figure 5.7: UKQCD Axial spatial t=0-14 matched unitary data sets

β	κ_{sea}	κ_{val}	am_{PS} (2-exp)	am_{PS} (MEM)
5.20	0.13500	0.13500	$0.405\ ^{+4}_{-5}$	$0.415\ ^{+15}_{-15}$
5.26	0.13450	0.13450	$0.509\ ^{+2}_{-2}$	$0.505\ ^{+11}_{-11}$
5.29	0.13400	0.13400	$0.577\ ^{+2}_{-2}$	$0.572\ ^{+5}_{-5}$
5.93	0.0	0.13390	$0.356\ ^{+1}_{-1}$	$0.375\ ^{+19}_{-19}$

Table 5.4: Comparison with UKQCD pseudoscalar masses

compare to. The error bars on the peaks are extremely large in this channel so very little statistical significance should be made of their shape.

Fig. 5.8 shows the results obtained for the pseudoscalar channel. Once again the only real trend seen is the ground state peak shifting towards lower energies as the mass of the quarks is decreased. The feature at around $\omega a \approx 1.5$ remains fairly steady. There is no sign of any dynamical quark effects in this channel. Table 5.4 is included to demonstrate that MEM is finding the same ground state as two-exponential fits.

Fig. 5.9 shows the results obtained for the Scalar channel. As with the axial spatial channel this is an extremely difficult channel to extract any information from using

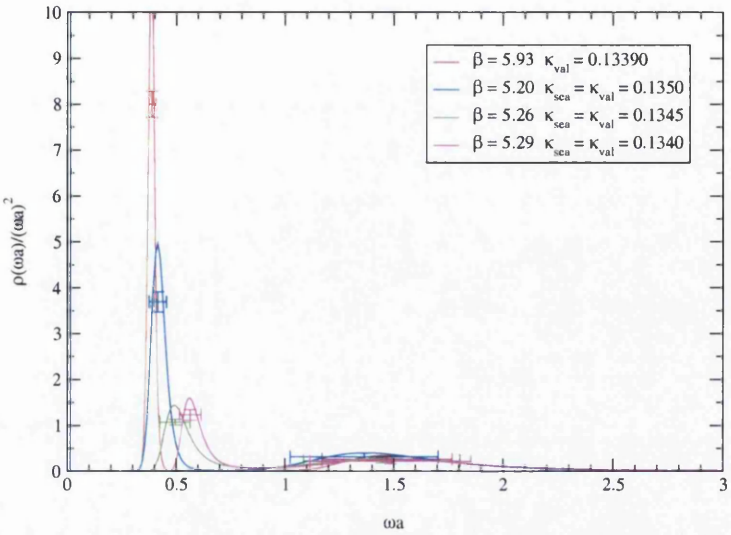

Figure 5.8: UKQCD Pseudoscalar t=0-14 matched unitary data sets

standard techniques, so there are no two-exponential values to compare to. The error bars on the peaks are large in this channel so very little statistical significance should be made of their shape. Again the differences between quenched and dynamical are slight. For a more detailed discussion on the scalar meson see § 5.4.

The Vector channel in Fig. 5.10 is where there may be signs of $\rho \rightarrow 2\pi$ decay. Note that the *pi*'s resulting from such a decay could not be at rest due to conservation of momentum, at rest the *rho* has angular momentum. The minimum momentum allowed for the 2 *pi* (one would have $\mathbf{p}$ and the other $\mathbf{p}$) would be

$$\mathbf{p}_{\text{min}} = \frac{\mathbf{2\pi}}{\mathbf{L}}\mathbf{a}^{-\mathbf{1}}. \tag{5.4}$$

Therefore the 2π feature is expected at $\omega = 2E_\pi$ where $E_\pi = \sqrt{m_\pi + \mathbf{p}^{\mathbf{2}}_{\text{min}}}$. Table 5.5 lists the value for $2E_\pi$ for each data set analysed.

Unfortunately there is no signal for ρ-decay in the current data set. Table 5.6 is a

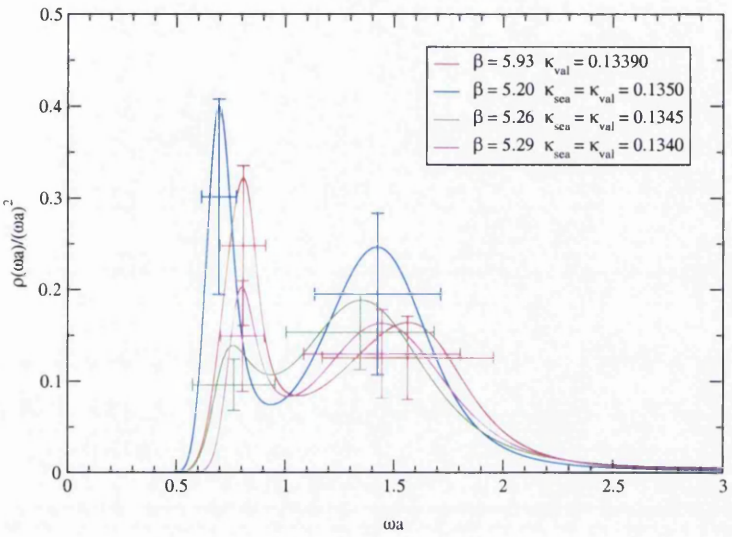

Figure 5.9: UKQCD Scalar t=0-14 matched unitary data sets

β	$\kappa_{\text{sea}} = \kappa_{\text{val}}$	$2E_\pi$
5.20	0.1358	0.89
5.20	0.1350	1.13
5.26	0.1345	1.29
5.29	0.1340	1.40

Table 5.5: Value of $2E_\pi$, the expected position for 2π feature indicating ρ-decay.

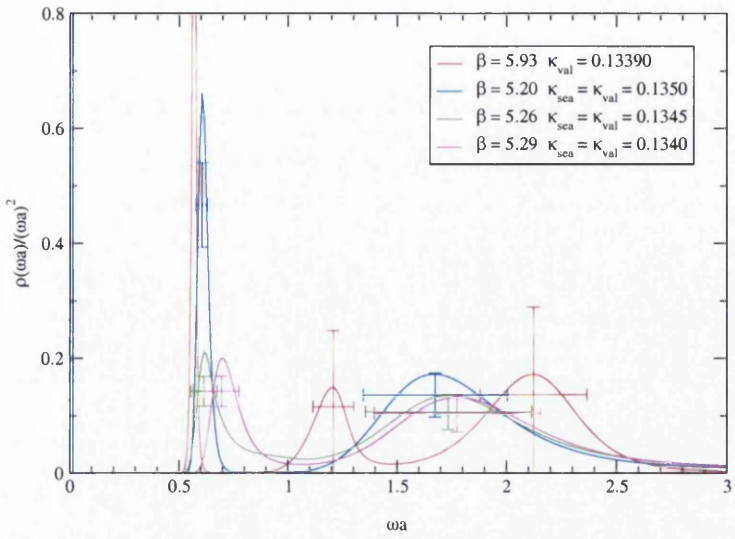

Figure 5.10: UKQCD Vector t=0-14 matched unitary data sets

β	κ_{sea}	κ_{val}	am_V (2-exp)	am_V (MEM)
5.20	0.13500	0.13500	$0.579\ ^{+7}_{-9}$	$0.610\ ^{+5}_{-5}$
5.26	0.13450	0.13450	$0.650\ ^{+4}_{-4}$	$0.637\ ^{+21}_{-21}$
5.29	0.13400	0.13400	$0.691\ ^{+6}_{-6}$	$0.719\ ^{+13}_{-13}$
5.93	0.0	0.13390	$0.563\ ^{+5}_{-5}$	$0.568\ ^{+5}_{-5}$

Table 5.6: Comparison with UKQCD vector masses

comparison of the Vector meson mass calculated from MEM and that from UKQCD two-exponential fits. The agreement between the two methods in much worse in this channel than for the pseudoscalar and axial temporal.

5.3.3 Lightest κ_{sea} data

Since no dynamical effects were seen within the matched unitary set, lets take a look at the results obtained from data at an even lighter κ_{sea} value. Almost certainly the lack of any observed dynamical effects in the spectral functions will be due to the still relatively heavy quarks simulated in the matched ensemble. While this data is not matched (the lattice spacing a is different) with all the results in § 5.3.2, due to the

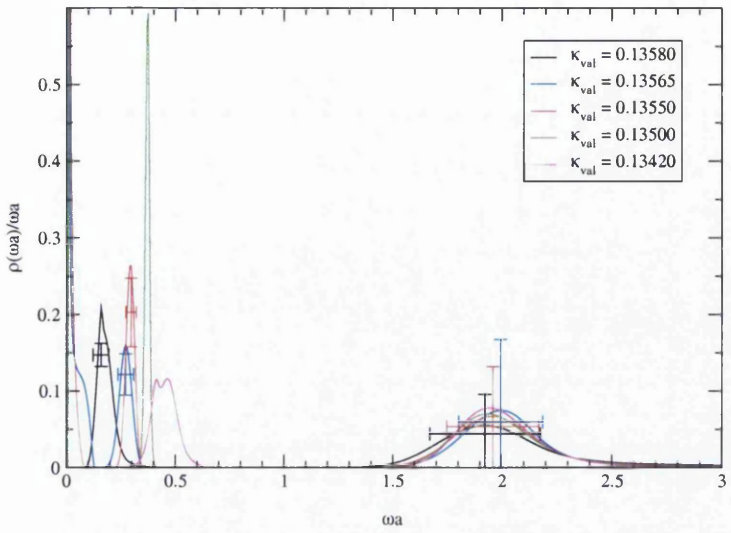

Figure 5.11: UKQCD Axial Temporal $\beta = 5.20$ $\kappa_{\text{sea}} = 0.1358$ t=0-14

restrictions in Eq. 5.1 for calculating c_{sw}, it is still interesting to analyse.

The graphs presented in this section display the spectral function calculated from each channel with varying valence quark mass. When the valence and sea quark mass values differ, it is known as a partially quenched approximation.

In the figures presented (5.11-5.15) here there are still no real signs of any differences due to the finite sea quark mass when making a comparison with those from the quenched data in § 5.3.1. The ground state peak tends to become lighter with the valence quark mass while the first excitation remains fairly steady, changing very little.

5.4 Non-singlet scalar meson

The study of the scalar meson (both singlet and non-singlet) is of great interest to particle physics. It is notoriously difficult to extract anything from this channel using

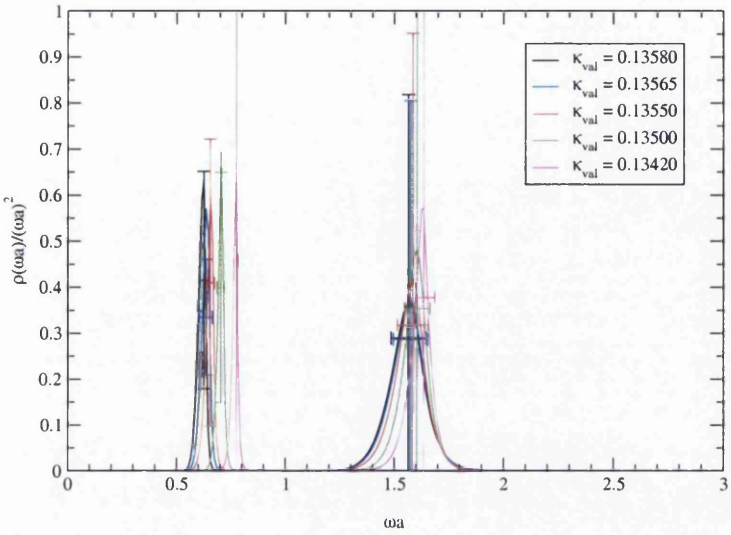

Figure 5.12: UKQCD Axial Spatial $\beta = 5.20$ $\kappa_{\text{sea}} = 0.1358$ t=0-14

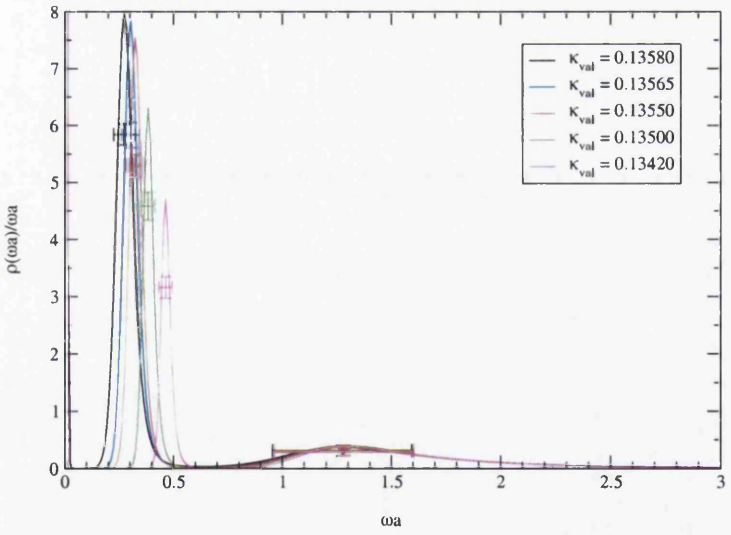

Figure 5.13: UKQCD Pseudoscalar $\beta = 5.20$ $\kappa_{\text{sea}} = 0.1358$ t=0-14

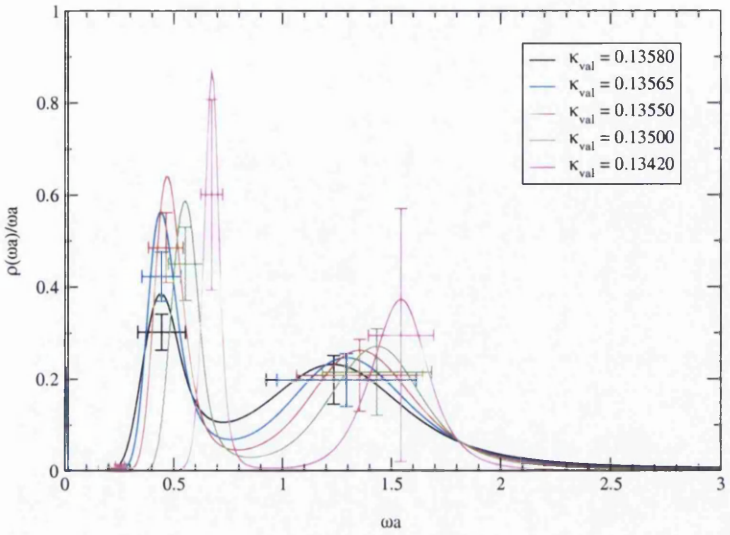

Figure 5.14: UKQCD Scalar $\beta = 5.20$ $\kappa_{\text{sea}} = 0.1358$ t=0-14

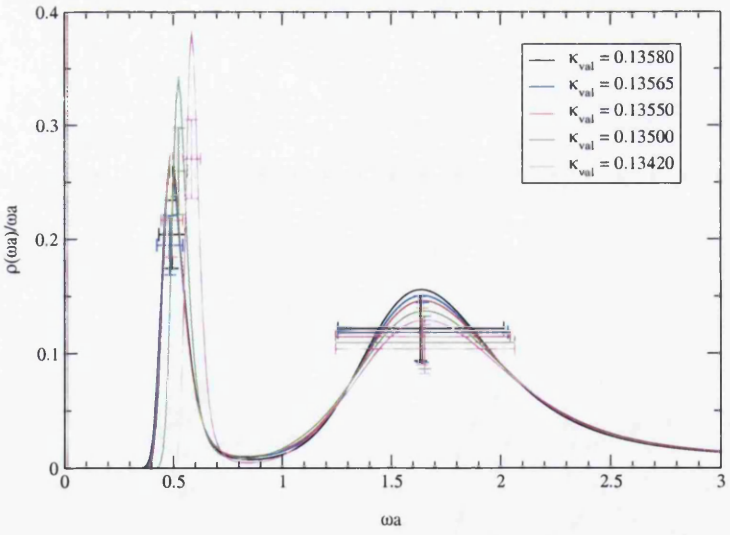

Figure 5.15: UKQCD Vector $\beta = 5.20$ $\kappa_{\text{sea}} = 0.1358$ t=0-14

conventional fitting procedures. Recently two "non" conventional approaches have been applied and have shed some light on both the singlet and non-singlet masses. The first approach [89] uses domain wall fermions and the second [90] (which analysed a similar data set to this work) used the Z_2 noise method. Both methods result in a mass of approximately $M_{a_0} = 1.0$ GeV in the chiral limit for the non-singlet scalar meson, though with relatively large errors.

In this thesis only the non-singlet scalar (a_0) will be studied and is the first time that results from conventional correlation functions have yielded the spectrum of the scalar meson in QCD.

Compared to the other channels studied the scalar channel's spectral function (Figs. 5.5, 5.9 and 5.14) has

(a) a very broad ground "state" and

(b) a second resonance of significant weight relative to the "ground state", which is not well separated from the ground "state".

This is presumably the reason why doing fits using sums of exponentials proves difficult.

In the quenched case the mass of the ground state M_{a_0} remains the same despite the change in quark mass (see Fig. 5.5). This is as Bardeen predicted in [91]. The matched unitary set doesn't display this pathology (i.e. M_{a_0} in the matched unitary set decreases sensibly as $m_q \to 0$).

The chiral extrapolation ($m_q \to 0$ or equivalently $aM_\pi^2 \to 0$ performed on the mass of the non-singlet scalar for the matched dynamical data sets only is displayed

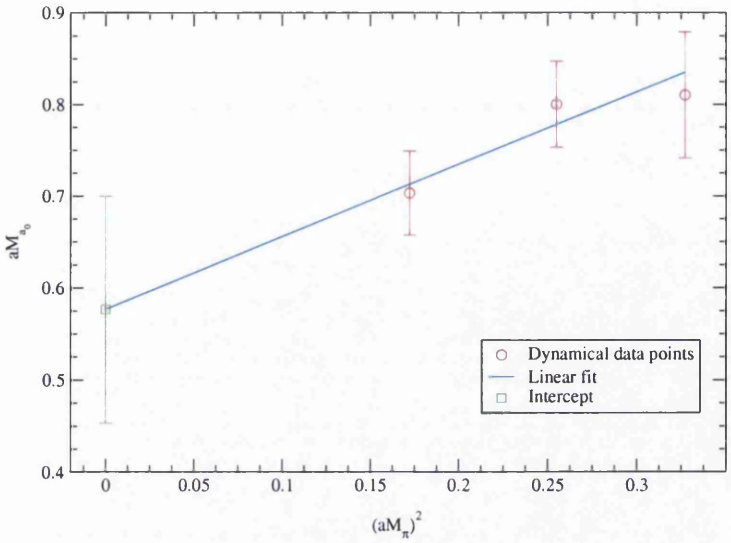

Figure 5.16: UKQCD Scalar mass chiral extrapolation

graphically in Fig. 5.16. This gives the non-singlet scalar mass calculated from MEM as $aM_{a_0} = 0.58(12)$ and multiplying by $a^{-1} = 1.92$ GeV (the errors on a^{-1} have been neglected since the 20% error on M_{a_0} will swamp it) gives $M_{a_0} = 1.1(2)$ GeV This agrees with the particle data books value of 985 MeV within the errors which are around 20%

5.5 Conclusions

- MEM has found ground state peaks in agreement with previous studies
- MEM has managed to extract information from local source and sink correlators in difficult channels such as the axial spatial and scalar.
- Unfortunately there have been no signs of dynamical effects in the channels analysed, even at the lightest quark mass.

- Non-singlet scalar channel was analysed and the prediction obtained for the mass of the a_0 is in agreement with experiment.

Chapter 6

Conclusions

Spectrum analysis techniques beyond the standard single- and multi-exponential fits, which assume a spectral density consisting of a series of isolated simple poles, will be required for lattice simulations beyond zero temperature and quenched QCD. In this thesis, the first attempt at applying the Maximum Entropy Method (MEM) to lattice data beyond the quenched approximation, but still at zero temperature, has been made.

In Chapter 3 MEM was successfully tested using known spectral function form. In Chapters 4 and 5 two field theories (a four fermion theory, the Gross-Neveu model in $d = 3$, referred to as GNM_3 and dynamical QCD) were studied, a summary of the results is now given.

6.1 Summary of the GNM_3 analysis (Chapter 4)

In the chirally broken phase:

- The elementary fermion f, the simplest $f\bar{f}$ bound state and the Goldstone boson π have all been shown to be sharp spectral features (i.e. simple poles).

- Estimates for the meson binding energy have been made for the first time.
- The additional contribution from disconnected fermion line diagrams cause the σ in the chirally broken phase to be more tightly bound than the conventional PS meson.

In the chirally symmetric phase:

- In the chirally symmetric phase, a broad resonance, whose features agree qualitatively with the large-N_f predictions, has been identified.
- Unfortunately there was no evidence of any feature at the two pion threshold in the σ channel, and therefore no evidence for $\sigma \to \pi\pi$ decay has been observed in this data.

6.2 Summary of QCD analysis (Chapter 5)

- MEM has found ground state peaks in agreement with previous studies.
- MEM has managed to extract information from difficult channels such as the axial spatial and scalar.
- Unfortunately there have been no signs of dynamical effects in the channels analysed, even at the lightest quark mass.
- Non-singlet scalar channel was analysed and a prediction obtained for the mass of the a_0 is in agreement with experiment.

6.3 Future work

6.3.1 GNM_3

It will be interesting to use MEM to study this model at non-zero temperature and/or density and perhaps looking at $d = 4$.

6.3.2 Dynamical QCD

No signs of the effects of introducing dynamical quarks have been seen in the five channels analysed from the UKQCD data set, even at the lightest quark mass. It will be interesting to apply the same analysis to any future data at lighter sea quark masses where the dynamical effects should be more prominent. Finite temperature QCD would be another area of particular interest for study with MEM, but this would require very large computer resources since finite temperature is acheived by decreasing the number of timeslices on the lattice. MEM requires a reasonable amount of data points so the only way to generate data for this would be to use and anisotropic lattice, i.e. a smaller spacing in the time direction than the three space directions.

Appendix A

Entropy definition

A.1 Monkey argument for entropy and prior probability

The prior probability can be written in the general form

$$P(f \in V) = \frac{1}{Z_S(\alpha)} \int_V [df] \Phi(\alpha S(f)), \tag{A.1}$$

where f is the image, V some domain, α an arbitrary constant and $Z_S(\alpha)$ is a normalisation constant. Φ is assumed to be a monotonic function of the entropy $S(f)$, therefore the most probable image f is obtained at the stationary point of $S(f)$.

The so-called "monkey argument", which is based on law of large numbers, can be used to determine the explicit forms for both Φ and $S(f)$.

First discretise the basis space x into N cells. Thus $f(x)$ also needs to be discretised as f_i. Now suppose a monkey throws M (assumed to be a large number) balls. The number of balls which land in the i^{th} cell is n_i and the probability that a ball lands in the same cells is p_i. The expectation value for the number of balls received

by the i^{th} cell λ_i is therefore given by

$$\lambda_i = Mp_i \quad \text{with} \quad \sum_{i=1}^{N} \lambda_i = M \tag{A.2}$$

The Poisson distribution (approximation for the Binomial distribution at large N and fixed λ_i) can be used to calculate the probability that the i^{th} cell receives n_i balls

$$P_{\lambda_i}(n_i) = \frac{\lambda_i^{n_i} e^{-\lambda_i}}{n_i!}. \tag{A.3}$$

Hence the probability that an entire distribution $\overrightarrow{n} = (n_1, \ldots, n_N)$ is realised is given by the product

$$P_{\overrightarrow{\lambda}}(\overrightarrow{n}) = \prod_{i=1}^{N} P_{\lambda_i}(n_i) = \prod_{i=1}^{N} \frac{\lambda_i^{n_i} e^{-\lambda_i}}{n_i!}. \tag{A.4}$$

with the normalisation given by $\sum_{n_i=0}^{\infty} P_{\lambda_i}(n_i) = 1$ $(i = 1, \ldots, N)$.

M is large, therefore n_i can also be large, so a small "quantum" q is introduced and a finite image f_i and default model (i.e. the expectation values) are defined as

$$f_i = qn_i, \quad m_i = q\lambda_i. \tag{A.5}$$

The probability $P(f \in V)$ can now be written as

$$P(f \in V) = \sum_{\overrightarrow{n} \in V} P_{\overrightarrow{\lambda}}(\overrightarrow{n}) \simeq \int_V \frac{\prod_{i=1}^{N} df_i}{q^N} \prod_{i=1}^{N} \frac{\lambda_i^{n_i} e^{\lambda_i}}{n_i!} \simeq \int_V \prod_{i=1}^{N} \frac{df_i}{\sqrt{f_i}} \frac{e^{S(f)/q}}{(2\pi q)^{N/2}}, \tag{A.6}$$

where the fact that q is small has been used in the conversion of the sum to an integral and Stirling's formula $n! = \sqrt{2\pi n}e^{n\log n - n}$ has been used to obtain the last expression.

Comparing Eqs. A.1 and A.6

$$q = \alpha^{-1}, \quad [df] = \prod_{i=1}^{N} \frac{df_i}{\sqrt{f_i}}, \quad Z_S(\alpha) = \left(\frac{2\pi}{\alpha}\right)^{N/2}. \tag{A.7}$$

A.2 Axiomatic construction of entropy

If a positive semi-definite distribution $f(x)$ is more likely than the distribution $g(x)$, then the Shannon-Jaynes entropy $S(f)$ needs to obey

$$\leftrightarrow S(f) > S(g). \tag{A.8}$$

If there is an external constraint on $f(x)$, e.g. $C(f(x)) = 0$, then the most plausible image is given by

$$\delta_f[S(f) - \lambda C(f)] = 0 \tag{A.9}$$

with λ a Lagrange multiplier. The explicit form of S can be fixed by considering the following axioms

Axiom 1 : Locality

$S(f)$ is a local functional of $f(x)$ without derivatives, i.e. there is no correlation between the images at different x.

This leads to a form

$$S(f) = \int dx\, m(x)\, \theta(f(x), x), \tag{A.10}$$

where $m(x)$ is a positive definite function which defines the integration measure. θ is an arbitrary local function of $f(x)$ and x without derivatives acting on f.

Axiom 2 : Coordinate Invariance

$f(x)$ and $m(x)$ transform as scalar densities under the coordinate transformation $x' = x'(x)$, i.e. $f(x)dx = f'(x')dx'$ and $m(x)dx = m'(x')dx'$. S is a scalar.

These constraints allow only two invariants for constructing S from Eq. A.10: $m(x)dx = m'(x')dx'$ and $f(x)/m(x) = f'(x')/m'(x')$. Hence

$$S(f) = \int dx\, m(x)\, \phi\left(\frac{f(x)}{m(x)}\right). \tag{A.11}$$

Axiom 3 : System Independence

If x and y are two independent variables, the image $F(x, y)$ and integration measure $m(x, y)$ are written in product form

$$F(x, y) = f(x)g(y), \quad m(x, y) = m_f(x)m_g(y). \tag{A.12}$$

Furthermore, the first variation of $S(F)$ with respect to $F(x,y)$ leads to an additive form with some functions $\alpha(x)$ and $\beta(y)$;

$$\frac{\delta S(F)}{\delta F(x,y)} = \alpha(x) + \beta(y) \tag{A.13}$$

Using this axiom, the images $f(x)$ and $g(y)$ can be independently determined

$$S(F) = \int dx \int dy\, m(x,y)\, \phi\left(\frac{F(x,y)}{m(x,y)}\right) \tag{A.14}$$

Acting the derivative $\partial^2/\partial x \partial y$ on Eq. A.14

$$Z\frac{d^2\sigma(Z)}{dZ^2} + \frac{d\sigma(Z)}{dZ} = 0 \tag{A.15}$$

where $Z \equiv F(x,y)/m(x,y) = (f(x)/m_f(x))(g(y)/m_g(y))$ and $\sigma(Z) = d\phi(Z)/dZ$. The solution to this differential equation is

$$\sigma(Z) = c_1 \log Z - c_0, \tag{A.16}$$

which leads to

$$\phi(Z) = c_1 Z \log Z - (c_0 + c_1)Z. \tag{A.17}$$

Thus the entropy can now be written as

$$S(f) = \int dx\, m(x)\, \phi\left(\frac{f(x)}{m(x)}\right) = \int dx\, f(x) \left[c_1 log\left(\frac{f(x)}{m(x)}\right) - (c_0 + c_1)\right]. \quad \text{(A.18)}$$

Since $(\delta/\delta f)^2 S(f) = c_1/f$ and $f \geq 0$, c_1 dictates the curvature of S. Thus c_1 is chosen to be -1 to have S bounded from above and normalised.

Axiom 4 : Scaling

If there is no external constraint on $f(x)$, the initial measure is recovered after the variation, i.e. $f(x) = m(x)$.

The unconstrained solution to of $\delta S(f)/\delta f = 0$ is $f(x) = m(x)e^{c_0/c_1}$ so $c_0 = 0$ in Eq. A.18. It is also convenient to add a constant on to the entropy to make $S(f = m) = 0$. Thus the Shannon-Jaynes entropy is defined as

$$S(f) = \int dx \left[f(x) - m(x) - f(x) \log\left(\frac{f(x)}{m(x)}\right)\right]. \quad \text{(A.19)}$$

Appendix B

Singular value decomposition

In this appendix a proof of the singular value decomposition (SVD), used in § 2.5.2, of a general $m \times n$ matrix M is given. The singular values of a matrix M are defined as the square root of the eigenvalues of $M^\dagger M$, which by definition is an $n \times n$ Hermitian matrix with real, non-negative eigenvalues. The following definitions for the norm of a vector $x \in \mathbb{C}^n$ and the spectral norm of M are also required.

$$\|x\|_2 = \left[\sum_{i=1}^{n} |x_i|^2\right]^{1/2} \tag{B.1}$$

$$\|M\|_2 = \left[\text{maximum eigenvalue of } M^\dagger M\right]^{1/2} \tag{B.2}$$

$$= \left[\max\left(x^\dagger M^\dagger M x\right)\right]^{1/2} (x \in \mathbb{C}^n, \|x\|_2 = 1). \tag{B.3}$$

SVD Theorem:

Let M be an $m \times n$ $(m \geq n)$ matrix, U and V be $m \times m$ and $n \times n$ unitary matrices respectively. Then M can be decomposed to

$$M = UWV^\dagger, \tag{B.4}$$

where W is an $m \times n$ diagonal matrix whose elements are the singular values of M ordered from largest to smallest. i.e. $W = \text{diag}(\xi_1, \xi_2, \ldots, \xi_n)$ with $\xi_1 \geq \xi_2 \geq \ldots \geq \xi_n \geq 0$

Proof:

The maximum singular value of M ξ_1 has a vector x_1 associated with it which satisfies the relation $\xi_1^2 = (x_1^\dagger M^\dagger M x_1)$, and also a vector y_1 which satisfies $Mx_1 = \xi_1 y_1$. Now these vectors x_1 and y_1 can be combined with non-square matrices U_2 and V_2 such that U_1 and V_1 defined below are $m \times m$ and $n \times n$ unitary matrices respectively.

$$U_1 = (y_1, U_2), \quad V_1 = (x_1, V_2) \tag{B.5}$$

These matrices can now be used to transform M into M_1

$$M_1 \equiv U_1^\dagger M V_1 = \begin{pmatrix} y_1^\dagger \\ U_2^\dagger \end{pmatrix} P \begin{pmatrix} x_1 & V_2 \end{pmatrix} = \begin{pmatrix} \xi_1 & y_1^\dagger M V_2 \\ 0 & U_2^\dagger M V_2 \end{pmatrix} \equiv \begin{pmatrix} \xi_1 & z_1^\dagger \\ 0 & Q_2 \end{pmatrix}, \tag{B.6}$$

where $z_1 \in \mathbb{C}^{n-1}$ and Q_2 is an $(m-1) \times (n-1)$ matrix. Now show that z_1 is, in fact, a null vector

$$\begin{aligned} \xi_1^2 &= \|M\|_2^2 = \|M_1\|_2^2 = \max\left(x^\dagger M^\dagger M x\right) && \text{(B.7)} \\ &\geq \frac{1}{\xi_1^2 + \|z_1\|_2^2} \begin{pmatrix} \xi_1 & z_1^\dagger \end{pmatrix} M_1^\dagger M_1 \begin{pmatrix} \xi_1 \\ z_1 \end{pmatrix} && \text{(B.8)} \\ &= \frac{1}{\xi_1^2 + \|z_1\|_2^2} \left[(\xi_1^2 + \|z_1\|_2^2)^2 + \|Q_2 z_1\|_2^2\right] && \text{(B.9)} \\ &\geq \xi_1^2 + \|z_1\|_2^2 && \text{(B.10)} \end{aligned}$$

This same procedure can now be applied to $Q_2, Q_3, \ldots$, using ξ_2 as the maximum singular value for Q_2 ($\xi_1 \geq \xi_2$) and so on. Thus

$$M_n = U^\dagger M V = \begin{pmatrix} \xi_1 & & \\ & \ddots & \\ & & \xi_n \\ 0 & \cdots & 0 \end{pmatrix} \equiv W \tag{B.11}$$

Thus the SVD $M = UWV^\dagger$ is proved.

The irrelevant components of U and W may be neglected so they become $m \times n$ and $n \times n$ matrices respectively. In this case U satisfies the condition $UU^\dagger = 1$, while $VV^\dagger = V^\dagger V = 1$.

Bibliography

[1] K. G. Wilson, Phys. Rev. D **10** (1974), 2445.

[2] H. Hamber and G. Parisi, Phys. Rev. Lett. **47** (1981), 1782.

[3] D. Weingarten, Phys. Lett. **109B** (1982), 57.

[4] S. Aoki *et. al.*, Phys. Rev. Lett. **84** (2000), 238.

[5] J. Garden, *Light Hadron Spectroscopy in Lattice QCD*, Ph.D. Thesis, University of Edinburgh, 2000.

[6] H. J. Rothe, *Lattice Gauge Theories - An Introduction*, World Scientific, Singapore, 1992.

[7] I. Montvay and G. Münster, *Quantum Fields on a Lattice*, Cambridge University Press, Cambridge, 1994.

[8] R. Gupta, *Introduction to Lattice QCD*, hep-lat/9807028, 1998.

[9] M. Lüsher, *Advanced Lattice QCD*, hep-lat/9802029, 1998.

[10] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965.

[11] (UKQCD Collaboration) C. R. Allton, (2001), hep-lat/0107021.

[12] S. Aoki *et. al.*, Phys. Rev. D **60** (1999), 114508.

[13] Y. Iwasaki, Tech. Report UTHEP-118, University of Tsukuba, 1983.

[14] K. G. Wilson, *New Phenomena in Subnuclear Physics*, Plenum, New York, 1975.

[15] B. Scheikholeslami and R. Wohlert, Nucl. Phys. B **185** (1985), 572.

[16] M. C. Chu, J. M. Grandy, S. Huang, and J. W. Negele, Phys. Rev. D **48** (1993), 3340.

[17] D. B. Leinweber, Phys. Rev. D **51** (1995), 6369.

[18] D. Makovoz and G. A. Miller, Nucl. Phys. B **468** (1996), 293.

[19] C. R. Allton and S. Capitani, Nucl. Phys.B **526** (1998), 463.

[20] T. Hashimoto, A. Nakamura, and I. O. Stamatescu, Nucl. Phys. B **400** (1993), 267.

[21] Ph.de Forcrand *et. al.* (QCD-Taro Collaboration), (2000), hep-lat/0008005.

[22] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, Univ. of Illinois Press, Urbana, 1949.

[23] E. T. Jaynes, Phys. Rev. **106** (1957), 620.

[24] B. R. Frieden, J. Opt. Soc. Am. **62** (1972), 511.

[25] M. Jarrell and J. E. Gubernatis, *Bayesian Inference and the Analytical Continuation of Imaginary-time Quantum Monte Carlo Data*, Phys. Rep. **269** (1996), 133–195.

[26] R. N. Silver *et. al*, Phys. Rev. Lett. **65** (1990), 496.

[27] R. N. Silver *et. al*, Phys. Rev. B **41** (1990), 2380.

[28] J. E. Gubernatis *et. al*, Phys. Rev. B **44** (1991), 6011.

[29] R. Preuss, W. Hanke, and W. von der Linden, Phys. Rev. Lett. **75** (1995), 1344.

[30] W. von der Linden, R. Preuss, and W. Hanke, J. Phys. **8** (1996), 3881.

[31] R. Preuss, W. Hanke, C. Grber, and H. G. Evertz, Phys. Rev. Lett. **79** (1997), 1122.

[32] S. E. Koonin, D. J. Dean, and K. Langanke, Phys. Rep. **278** (1997), 1.

[33] N. Wu, *The Maximum Entropy Method*, Springer-Verlag, Berlin, 1997.

[34] H. Jeffreys, *Theory of Probability*, Third ed., Oxford Univ. Press, Oxford, 1998.

[35] R. K. Bryan, *Maximum entropy analysis of oversampled data problems*, Eur. Biophys. J. **18** (1990), 165–174.

[36] J. Skilling, *Maximum Entropy and Bayesian Methods*, pp. 45–52, Kluwer, Academic Publisher, London, 1989.

[37] S. F. Gull and G. J. Daniel, Nature **272** (1978), 686.

[38] E. T. Jaynes, *Maximum Entropy and Bayesian Methods in Applied Statistics*, pp. 26–58, Cambridge Univ. Press, Cambridge, 1986.

[39] J. Skilling, *Maximum Entropy and Bayesian Methods in Science and Engineering*, vol. 1, pp. 173–187, Kluwer, Academic Publisher, London, 1988.

[40] F. Chatelin, *Eigenvalues of Matrices*, Wiley & Sons, New York, 1993.

[41] W. H. Press *et. al.*, *Numerical recipes*, Second ed., Cambridge Univ. Press, Cambridge, 1994.

[42] Y. Nakahara, M. Asakawa, and T. Hatsuda, *Maximum entropy analysis of the spectral functions in lattice qcd*, **47** (2000).

[43] Z. Fodor and S. D. Katz, Phys. Lett. B. **534** (2002), 87.

[44] C. R. Allton *et. al.*, (2002), hep-lat/0204010.

[45] C. Bernard, T. A. DeGrand, C. DeTar, S. Gottlieb, A. Krasnitz, M. C. Ogilvie, R. L. Sugar, and D. Toussaint, Phys. Rev. Lett. **68** (1992), 277.

[46] T. H. Hansson and I. Zhed, Nucl. Phys. B **374** (1992), 277.

[47] I. M. Barbour, Nucl. Phys. B (Proc. Suppl.) **26** (1992), 22.

[48] E. Fermi, Z. Physik 88 **161** (1934).

[49] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108** (1957), 1175.

[50] Y. Nambu and G. Jona-Lasinio, Phys. Rev. **122** (1991), 345.

[51] D. Ebert and H. Reinhardt, Nucl. Phys. B **271** (1986), 188.

[52] A. Dhar and S. R. Wadia, Phys. Rev. Lett. **52** (1984), 959.

[53] T. Hatsuda and T. Kunihiro, Phys. Rep. **247** (1994), 221.

[54] B. Holdom, Phys. Lett. B **150** (1985), 301.

[55] K. Yamawaki, M. Bando, and K. Matumoto, Phys. Rev. Lett. **56** (1986), 1335.

[56] Y. Nambu, New Trends in Physics, Proceedings of the XI International Symposium on Elementary Particle Physics (Kazimeirz, Poland, 1988) (Z. Adjuk, S. Pokorski, and A. Trautman, eds.), World Scientific, 1989.

[57] W. A. Bardeen, C. T. Hill, and M. Lindner, Phys. Rev. D **41** (1990), 1647.

[58] S. J. Hands and J. B. Kogut, Nucl. Phys. B **529** (1998), 383.

[59] M. Asakawa and K. Yazaki, Nucl. Phys. A **504** (1989), 668.

[60] S. P. Klevansky, Rev. Mod. Phys. **64** (1992), 649.

[61] R. F. Dashen and B. Hasslacher, Phys. Rev. D **12** (1975), 2443.

[62] U. Wolff, Phys. Lett. B **157** (1985), 303.

[63] F. Karsch, J. Kogut, and H. W. Wyld, Nucl. Phys. B **280** (1987), 289.

[64] B. Rosenstein, B. J. Warr, and S. H. Park, Phys. Rep. **205** (1991), 59.

[65] S. J. Hands, A. Kocic, and J. B. Kogut, Ann. Phys. **224** (1993), 29.

[66] S. J. Hands, A. Kocic, and J. B. Kogut, Nucl. Phys. B **390** (1993), 355.

[67] S. K. Ma, *Modern Theory of Critical Phenomenon*, Benjamin, 1976.

[68] J. A. Gracey, Int. J. Mod. Phys. A **6** (1991), 395.

[69] J. A. Gracey, A **9** (1994), 567.

[70] J. A. Gracey, Phys. Lett. B **308** (1994), 65.

[71] J. A. Gracey, Phys. Lett. D **50** (1994), 2840.

[72] L. Karkkainen, R. Lacaze, P. Lacock, and B. Peterson, Nucl. Phys. B **415** (1994), 781.

[73] A. Kocic and J. Kogut, Phys. Rev. Lett. **74** (1995), 3109.

[74] A. Kocic and J. Kogut, Nucl. Phys. B **455** (1995), 229.

[75] S. J. Hands, S. Kim, and J. B. Kogut, Nucl. Phys. B **442** (1995), 364.

[76] I. M. Barbour, S. J. Hands, J. B. Kogut, M. P. Lombardo, and S. E. Morrison, Nucl. Phys. B **557** (1999), 327.

[77] R. Gupta, G. Guralnik, G. W. Kilcup, and S. R. Sharpe, Phys. Rev. D **43** (1991), 2003.

[78] S. J. Hands, J. B. Kogut, and C. G. Strouthos, Phys. Rev. D **65** (2002), 114507.

[79] T. Hatsuda, (2002), hep-lat/0208059.

[80] N. Ishizuka and T. Yamazaki, (2002), hep-lat/0209058.

[81] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Phys. Lett. B **195** (1987), 216.

[82] Z. Sroczynski, Ph.D. Thesis, University of Edinburgh, 1998.

[83] S. M. Pickles, Ph.D. Thesis, University of Edinburgh, 1998.

[84] (UKQCD Collaboration) C. R. Allton, Phys. Rev. D **60** (1999), 034507.

[85] B. Joó, Ph.D. Thesis, University of Edinburgh, 2000.

[86] C. R. Allton *et. al.*, Nucl. Phys. B **407** (1993), 331.

[87] K. Jansen and R. Sommer, Nucl. Phys. B **530** (1998), 185.

[88] R. G. Edwards, U. M. Heller, and T. R. Klassen, Phys. Rev. Lett. **80** (1998), 3448.

[89] S. Prelovsek and K. Orginos, (2002), hep-lat/0209132.

[90] (UKQCD Collaboration) A. Hart, C. McNeile, and C. Michael, (2002), hep-lat/0209063.

[91] W. Bardeen *et. al.*, Phys. Rev. D. **65** (2002), 014509.