Complex Langevin dynamics as a cure for the sign problem.

James, Frank

Complex Langevin dynamics as a cure for the sign problem

Frank James
Swansea University 2012

Submitted to Swansea University in fulfilment of the requirements for the Degree of Doctor of Philosophy

Declaration

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

Signed ..... (Frank James)

Date 11/6/12

Statement 1

The thesis is the results of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s). Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed ..... (Frank James)

Date 11/6/12

Statement 2

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

Signed ..... (Frank James)

Date 11/6/12

Contents

List of Figures

List of Tables

List of publications

- Adaptive stepsize and instabilities in complex Langevin dynamics,
  G. Aarts, F. A. James, E. Seiler and I. -O. Stamatescu, Phys.Lett. B687 (2010) 154-159, arXiv:0912.0617 [hep-lat]
- On the convergence of complex Langevin dynamics: the three-dimensional XY model at finite chemical potential,
  G. Aarts and F. A. James, JHEP 1008 (2010) 020, arXiv:1005.3468 [hep-lat]
- Complex Langevin: etiology and diagnostics of its main problem,
  G. Aarts, F. A. James, E. Seiler and I. -O. Stamatescu, Eur.Phys.J. C71 (2011) 1756, arXiv:1101.3270 [hep-lat]
- Complex Langevin dynamics in the SU(3) spin model at nonzero chemical potential revisited,
  G. Aarts and F. A. James, JHEP 1201 (2012) 118, arXiv:1112.4655 [hep-lat]

Proceedings

- The XY model at finite chemical potential using complex Langevin dynamics,
  G. Aarts and F. A. James, PoS LATTICE2010 (2010) 321, arXiv:1009.5838 [hep-lat]

- Complex Langevin dynamics: criteria for correctness,
  G. Aarts, F. A. James, E. Seiler and I. -O. Stamatescu, PoS LATTICE2011 (2011) 197, arXiv:1110.5749 [hep-lat]

Acknowledgements

I would like to thank Gert Aarts for constant and continued support throughout, Erhard Seiler and Nucu Stamatescu for numerous discussions and also funding from STFC, Hadron Physics II, Strongnet, DiRAC Facility and the Large Facilities Capital Fund of BIS.

Abstract

Theories with a chemical potential are difficult to treat numerically because the action is complex and therefore methods based on a probability interpretation of the weight break down. This is an issue known as the sign problem. Complex Langevin dynamics was first proposed in the early 1980s and does not rely in a probability interpretation of the weight, so it can in principle be applied even where there is a severe sign problem. However, the combined problems of numerical instabilities and incorrect convergence impeded such early studies. In this work, the problem of runaway trajectories is cured by the use of a general adaptive stepsize procedure, which can be applied to both abelian and non-abelian theories. A study of the three-dimensional XY model at non-zero chemical potential follows, in which the problem of incorrect convergence is encountered. A formal justification of complex Langevin dynamics is given, from which a set of criteria are derived which can be used to test the validity of results. These ideas are applied to the SU(3) spin model, which is found to pass them all and therefore give correct results. An improved integration algorithm, which eliminates leading order step size corrections, is outlined and shown to give improved results.

Chapter 1

Introduction

Field theories with a complex action are difficult to treat nonperturbatively, because the weight in the partition function is not real. Standard numerical approaches based on a probability interpretation and importance sampling will then typically break down, which is commonly referred to as the sign problem. This is a pressing problem for QCD, where a nonperturbative determination of the phase diagram in the plane of temperature and baryon chemical potential is still lacking [1].

Several methods have been developed to explore at least part of the phase diagram [2–12], but in general these can only be applied in a limited region. Recent years have also seen an intense study of the sign problem in QCD and related theories, which has led to new formulations [13, 14] and considerable insight into how the complexity of the weight interplays with physical observables [15–24]. A constrained sampling of field space, yielding a joint probability distribution for only a small number of observables, is the notion behind the factorization/density of states/histogram approaches [12, 25–27]. In some theories, it is possible to group degrees of freedom together in such a way that the sign problem is manifestly absent [28–31]. This is the idea behind the meron cluster algorithm [32] and it has been applied recently to random matrix theory at finite chemical potential [33]. Recent success-

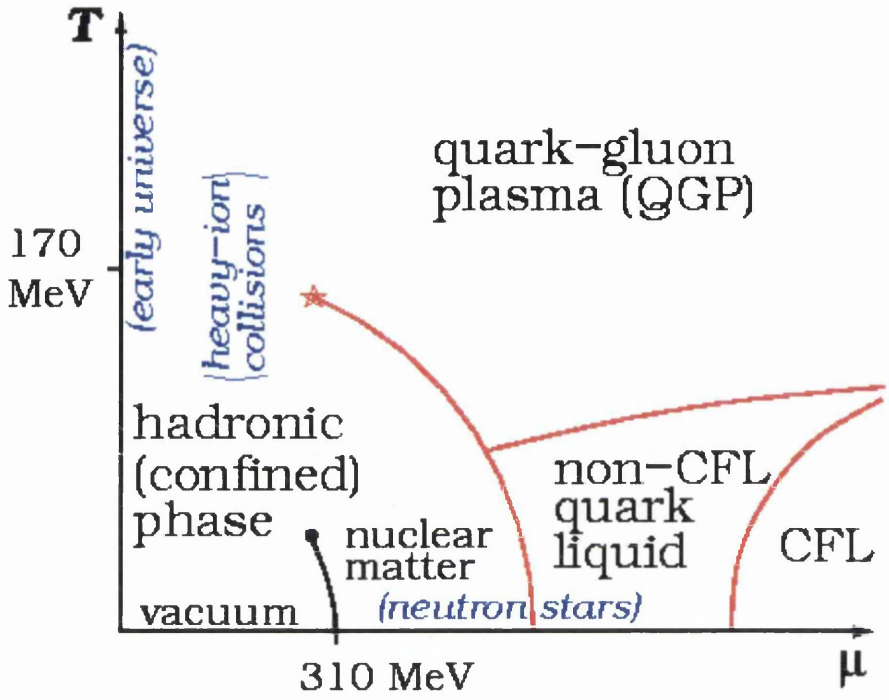


Figure 1.1: Conjectured phase diagram

ful applications of this have been to models derived from QCD in combined strong-coupling and heavy-quark expansions [34–40]. This suggests that the sign problem is not a problem of principle for a theory, but instead tied to the formulation and/or algorithm. For QCD an exact reformulation without a sign problem has unfortunately not (yet) been found.

As a result, although constructing the phase diagram of QCD from first principles remains out of reach of current techniques [41], it is nevertheless possible to derive the general phase structure from other arguments. A conjectured QCD phase diagram as a function of quark chemical potential [1] is shown in Figure 1.1. The region of small chemical potentials along the T-axis can be studied by various methods, because here the sign problem is not severe and so the techniques based on standard approaches can be applied (see

below). Heavy ion collisions, which probe the Quark-gluon plasma phase are of particular interest here. In the very high density regime, one can use the fact that the system is weakly coupled, yielding the colour-flavour locking phase (CFL). Unfortunately, this regime is only valid at exponentially high densities and so is unlikely to be of relevance to either astrophysics or experiments. This leaves the majority of the phase diagram without a reliable tool to determine it. Unfortunately the region of serious interest in astrophysics, that of cold dense matter, lies in this regime.

The most simple way of removing the complex phase from the weight is to absorb it into the observable. The total observable is then found from the ratio of two expectation values sampled with respect to the *phase quenched* weight,

$$\langle O \rangle = \frac{\int D\phi\, O(\phi) e^{i\varphi} \left| e^{-S(\phi)} \right|}{\int D\phi\, e^{i\varphi} \left| e^{-S(\phi)} \right|}, \tag{1.1}$$

$$= \frac{\langle O e^{i\varphi} \rangle_{\mathrm{pq}}}{\langle e^{i\varphi} \rangle_{\mathrm{pq}}}. \tag{1.2}$$

The phase quenched weight is real and positive, therefore standard techniques employing importance sampling can be used. Unfortunately, this process breaks down as the volume Ω is taken to infinity. This is because the denominator is a ratio of the partition functions which vanishes exponentially in the thermodynamic limit,

$$\langle e^{i\varphi} \rangle_{\mathrm{pq}} = \frac{Z_{\mathrm{full}}}{Z_{\mathrm{pq}}} = e^{-\Omega \Delta f}, \tag{1.3}$$

where Δf is the difference in free energies between the phase quenched and full theories. Therefore the simulation is required to run exponentially longer in order to average out the fluctuations, nullifying the gain from simulating a real and positive weight. However, the method can be used to great effect when the severeness of the sign problem, as measured by $\langle e^{i\varphi} \rangle_{\mathrm{pq}}$ is not too

close to zero, which limits use to small volumes and chemical potentials.

Although a sign problem may be present with a real chemical potential, taking an imaginary chemical potential $\mu = i\mu_I$ makes the action real. By expanding in μ^2, it can be seen that observables which are even in μ will be continuous across the boundary between real and imaginary chemical potential at $\mu^2 = 0$. Therefore, it is possible to perform simulations with a real action and extrapolate to regions of small real chemical potential by analytic continuation. As with any technique involving and expansion, the reliability of results breaks down for larger chemical potentials.

An alternative way of examining the region of small chemical potential is to perform a Taylor expansion in μ. The coefficients of the expansion will involve expectation values of observables at $\mu = 0$, where there is no sign problem and can therefore be computed using standard methods. Provided the chemical potential is small the Taylor expansion method allows the $\mu \neq 0$ region of the phase diagram to be probed.

1.1 Stochastic quantization

The stochastic quantization method provides an alternative way of defining quantum expectation values compared to the canonical or path integral methods. The method begins by first observing the important analogies between Euclidean quantum field theory and classical statistical mechanics, specifically that the measure of a quantum path integral is closely related to the Boltzmann distribution for a classical statistical system. This means that the quantum path integral

$$\langle O(\phi) \rangle = \frac{\int D\phi \, O(\phi) e^{-\frac{1}{\hbar} S(\phi)}}{\int D\phi \, e^{-\frac{1}{\hbar} S(\phi)}} \tag{1.4}$$

may, with the identification of $1/\hbar = kT$, be interpreted as a statistical expectation value with respect to a system in equilibrium at a temperature

T.

The fundamental principle of stochastic quantization is to consider the measure of the Euclidean path integral to be the equilibrium distribution of a stochastic process. The task is then to define a process which has this as its stationary distribution. Brownian motion provides a simple example of such a process that yields a correct distribution in the limit of large times. Start by writing the equation of motion for a particle in a 3d viscous fluid under the influence of random noise kicks,

$$m\dot{v}_i = -\alpha v_i + \eta_i, \tag{1.5}$$

with η_i random noise representing interactions with other particles and α the coefficient of friction. The stochastic sources η_i represent Gaussian noise with width λ,

$$\langle \eta_i(t)\eta_j(t')\rangle = 2\lambda\delta(t-t')\delta_{ij}, \qquad \langle \eta_i(t)\rangle = 0. \tag{1.6}$$

The equation of motion can be solved for time t, giving

$$v_i(t) = e^{-\frac{\alpha}{m}t}v_i(0) + \frac{1}{m}\int_0^t e^{-\frac{\alpha}{m}(t-\tau)}\eta_i(\tau)d\tau. \tag{1.7}$$

Assuming for simplicity that $v_i(0) = 0$, the average kinetic energy is given by

$$\begin{aligned}\frac{1}{2}m\langle v_i(t)v_i(t)\rangle &= \frac{1}{2}m\frac{1}{m^2}\int_0^t\int_0^t e^{-\frac{\alpha}{m}(2t-\tau-\tau')}\langle \eta_i(\tau)\eta_i(\tau')\rangle\, d\tau d\tau', \\ &= \frac{3}{2}\frac{\lambda}{\alpha}\left(1-e^{-2\frac{\alpha}{m}t}\right).\end{aligned} \tag{1.8}$$

After large times the particle comes into thermal equilibrium with the fluid, so that all t-dependence vanishes, and therefore should have a kinetic energy

equal to $\frac{3}{2}kT$. This allows the width of the noise to be identified as

$$\lambda = \alpha kT. \tag{1.9}$$

This simple analysis shows that it is possible to derive results about bulk properties, such as average kinetic energy, from an analysis of the microscopic movements of individual particles by averaging over a suitably chosen stochastic source.

This idea of measuring with respect to the stationary distribution of a stochastic process can be generalised to field theory, which therefore provides an alternative way of defining expectation values.

The procedure begins by adding an additional fictional time-like parameter ϑ, the Langevin time, to the degrees of freedom,

$$\phi \to \phi(\vartheta). \tag{1.10}$$

The demand is then that ϕ evolves in Langevin time according to some stochastic process which relaxes to a unique stationary distribution. This stationary distribution should be the "correct" distribution, e^{-S}, so that averages over noise will be equivalent to averages with respect to this weight. A suitable stochastic differential equation is the Langevin equation,

$$\frac{\partial}{\partial\vartheta}\phi(\vartheta) = -\frac{\delta S}{\delta\phi(\vartheta)} + \eta(\vartheta), \tag{1.11}$$

with the stochastic noise term η obeying the relations

$$\langle\eta(\vartheta)\eta(\vartheta')\rangle = 2\delta(\vartheta - \vartheta'), \qquad \langle\eta(\vartheta)\rangle = 0, \tag{1.12}$$

i.e. η is Gaussian,

$$P(\eta) \sim e^{-\eta^2/4}. \tag{1.13}$$

To ascertain that the Langevin equation does indeed generate a sequence of

correctly distributed configurations, it is sufficient to show that the process is reversible. This property, also known as detailed balance, requires that the forward and reverse transition probabilities $P_{\rm tr}(\phi \to \phi')$ between two states ϕ and ϕ', are balanced with the probability of being in a particular state, $P_{\rm eq}(\phi)$,

$$P_{\rm eq}(\phi)P_{\rm tr}(\phi \to \phi') = P_{\rm eq}(\phi')P_{\rm tr}(\phi' \to \phi). \tag{1.14}$$

The desired probability distribution is $P_{\rm eq}(\phi) \sim e^{-S(\phi)}$.

In order to see that the Langevin process satisfies this, we start by discretizing time $\vartheta = n\epsilon$, and expanding to lowest order in ϵ. This gives the update rule

$$\phi(n+1) = \phi(n) - \epsilon\frac{\delta S}{\delta\phi(n)} + \sqrt{\epsilon}\eta(n). \tag{1.15}$$

The probability to transition from state ϕ to another state ϕ' after a single update can then be found by reversing the update rule Eq. (1.15),

$$P_{\rm tr}(\phi \to \phi') = N_0 \exp\left\{-\frac{1}{4}\left[\frac{\phi'-\phi}{\sqrt{\epsilon}} + \sqrt{\epsilon}\frac{\delta S}{\delta\phi}\right]^2\right\}. \tag{1.16}$$

From Eq. (1.16) it follows that

$$\frac{P_{\rm tr}(\phi \to \phi')}{P_{\rm tr}(\phi' \to \phi)} = \exp\left\{-\frac{1}{2}\left((\phi'-\phi)\frac{\delta S}{\delta\phi} - (\phi-\phi')\frac{\delta S}{\delta\phi'}\right) + O(\epsilon)\right\}. \tag{1.17}$$

Assuming the difference $\phi' - \phi$ is small, the action can be expanded to lowest order around ϕ, giving

$$S(\phi') = S\Big(\phi + (\phi'-\phi)\Big) = S(\phi) + (\phi'-\phi)\frac{\delta S}{\delta\phi} + O(\epsilon). \tag{1.18}$$

It then follows that

$$(\phi'-\phi)\frac{\delta S}{\delta\phi} = S(\phi') - S(\phi) + O(\epsilon), \tag{1.19}$$

and therefore that in the limit of $\epsilon \to 0$, the ratio of transition probabilities is

$$\frac{P_{\mathrm{tr}}(\phi \to \phi')}{P_{\mathrm{tr}}(\phi' \to \phi)} = e^{S(\phi)-S(\phi')}, \tag{1.20}$$

which is precisely the requirement of detailed balance in Eq. (1.14). At this stage it is important to note that this justification of Langevin dynamics is reliant upon the action being real; if this were not the case then it would not be possible to interpret the equilibrium distribution e^{-S} as a probability. There exist other more formal proofs of the Langevin method involving the dual Fokker-Planck equation, but again they also rely on the action being real.

1.2 Complex Langevin dynamics

Although the formal proofs for the correctness of the Langevin equation break down when the action is complex, it is nevertheless possible to write a complex Langevin equation. Despite the lack of a formal proof, complex Langevin dynamics [42,43] offers the possibility of a general solution to the sign problem. In this formulation the fields, denoted here collectively as ϕ, are supplemented with an additional fictional Langevin time ϑ and the system evolves according to the stochastic equation,

$$\frac{\partial \phi_x(\vartheta)}{\partial \vartheta} = -\frac{\delta S[\phi;\vartheta]}{\delta \phi_x(\vartheta)} + \eta_x(\vartheta). \tag{1.21}$$

In this case, the action is complex and as a result the fields are *complexified* to

$$\phi \to \phi^{\mathrm{R}} + i\phi^{\mathrm{I}}. \tag{1.22}$$

The Langevin equations then read (using general complex noise)

$$\frac{\partial \phi_x^{\mathrm{R}}}{\partial \vartheta} = K_x^{\mathrm{R}} + \sqrt{N_{\mathrm{R}}}\eta_x^{\mathrm{R}}, \qquad K_x^{\mathrm{R}} = -\mathrm{Re}\frac{\delta S}{\delta \phi_x}\Big|_{\phi \to \phi^{\mathrm{R}} + i\phi^{\mathrm{I}}}, \tag{1.23a}$$

$$\frac{\partial \phi_x^{\mathrm{I}}}{\partial \vartheta} = K_x^{\mathrm{I}} + \sqrt{N_{\mathrm{I}}}\eta_x^{\mathrm{I}}, \qquad K_x^{\mathrm{I}} = -\mathrm{Im}\frac{\delta S}{\delta \phi_x}\Big|_{\phi \to \phi^{\mathrm{R}} + i\phi^{\mathrm{I}}}. \tag{1.23b}$$

The strength of the noise in the real and imaginary components of the Langevin equation is constrained via $N_{\mathrm{R}} - N_{\mathrm{I}} = 1$, and the noise furthermore satisfies

$$\langle \eta_x^{\mathrm{R}}(\vartheta) \rangle = \langle \eta_x^{\mathrm{I}}(\vartheta) \rangle = \langle \eta_x^{\mathrm{R}}(\vartheta) \eta_y^{\mathrm{I}}(\vartheta') \rangle = 0, \tag{1.24a}$$

$$\langle \eta_x^{\mathrm{R}}(\vartheta) \eta_y^{\mathrm{R}}(\vartheta') \rangle = \langle \eta_x^{\mathrm{I}}(\vartheta) \eta_y^{\mathrm{I}}(\vartheta') \rangle = 2\delta(x-y)\delta(\vartheta - \vartheta'), \tag{1.24b}$$

i.e., it is Gaussian. Since the complex action is only used to generate the drift terms but not for importance sampling, complex Langevin dynamics can potentially avoid the sign problem.[1]

In the limit of infinite Langevin time, noise averages of observables should equal the standard quantum expectation values. When the action is real, formal proofs that Langevin dynamics converges to the correct distribution can be formulated, see for instance the justification based on detailed balance of Sec. 1.1 and arguments that use properties of the associated Fokker-Planck equation [50]. If the action is complex and the Langevin dynamics extends into the expanded complexified space, these proofs no longer hold. Nevertheless, a formal derivation of the validity of the approach can still be given, employing holomorphicity and the Cauchy-Riemann equations. We sketch here the basic notion, suppressing all indices for notational simplicity, and refer to Chapter 5 and Ref. [58] for details.

Associated with the Langevin process (1.23) is a (real and positive) probability density $P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}; \vartheta]$, which evolves according the Fokker-Planck equa-

[1] Early studies of complex Langevin dynamics can be found in, e.g., Refs. [44–49]. Ref. [50] contains a further guide to the literature. More recent work includes Refs. [51–60].

tion

$$\frac{\partial P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}; \vartheta]}{\partial \vartheta} = L^T P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}; \vartheta], \tag{1.25}$$

with the Fokker-Planck operator

$$L^T = \frac{\partial}{\partial \phi^{\mathrm{R}}} \left[N_{\mathrm{R}} \frac{\partial}{\partial \phi^{\mathrm{R}}} - K^{\mathrm{R}} \right] + \frac{\partial}{\partial \phi^{\mathrm{I}}} \left[N_{\mathrm{I}} \frac{\partial}{\partial \phi^{\mathrm{I}}} - K^{\mathrm{I}} \right]. \tag{1.26}$$

Stationary solutions of this Fokker-Planck equation are only known in very special cases [46, 56, 58, 61]. Expectation values obtained by solving the stochastic process should then equal

$$\langle O \rangle_{P(\vartheta)} = \frac{\int D\phi^{\mathrm{R}} D\phi^{\mathrm{I}} \, P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}; \vartheta] O[\phi^{\mathrm{R}} + i\phi^{\mathrm{I}}]}{\int D\phi^{\mathrm{R}} D\phi^{\mathrm{I}} \, P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}; \vartheta]}. \tag{1.27}$$

However, we may also consider expectation values with respect to a complex weight $\rho[\phi; \vartheta]$,

$$\langle O \rangle_{\rho(\vartheta)} = \frac{\int D\phi \, \rho[\phi; \vartheta] O[\phi]}{\int D\phi \, \rho[\phi; \vartheta]}, \tag{1.28}$$

where, using Eq. (1.21), ρ evolves according to a complex Fokker-Planck equation

$$\frac{\partial \rho[\phi; \vartheta]}{\partial \vartheta} = L_0^T \rho[\phi; \vartheta], \qquad L_0^T = \frac{\partial}{\partial \phi} \left[\frac{\partial}{\partial \phi} + \frac{\partial S}{\partial \phi} \right]. \tag{1.29}$$

This equation has the desired stationary solution $\rho[\phi] \sim \exp(-S)$.

Under some assumptions and relying on holomorphicity and partial integration [58], one can show that these expectation values are equal, and

$$\langle O \rangle_{P(\vartheta)} = \langle O \rangle_{\rho(\vartheta)}. \tag{1.30}$$

If it can subsequently be shown that

$$\lim_{\vartheta \to \infty} \langle O \rangle_{\rho(\vartheta)} = \langle O \rangle_{\rho(\infty)}, \qquad \rho(\phi; \infty) \sim \exp(-S), \tag{1.31}$$

the applicability of complex Langevin dynamics is demonstrated. These ar-

guments are explained in detail in Chapter 5 where it is shown that indeed this is the case, provided a certain set of criteria are satisfied.

1.3 Thesis outline

The complexification of the fields introduces new degrees of freedom, which are typically unbounded and can potentially follow divergent trajectories which renders numerical simulations unstable. Indeed, in most cases the classical flow will have unstable fixed points, meaning that a classical solution would be driven out to infinity in a finite time. However, the introduction of the stochastic noise term has the effect of kicking off these trajectories and therefore keeping the dynamics stable.

Nevertheless, in most theories, the force term grows very quickly (often exponentially) in the direction of the imaginary part of the complexified variables. This means that if the field is sent out on a long trajectory by the classical flow, it can become difficult to do the numerical integration correctly. Since this problem is numerical in origin, it can be overcome by careful integration when the force terms become large. An algorithm for controlling the force term and adapting the stepsize appropriately is explained and tested in detail in Chapter 2.

Following from the analysis of the adaptive stepsize algorithm, a detailed study of the XY model at nonzero chemical potential follows in Chapter 3. This model is useful as a test theory because it can be rewritten in terms of dual variables which makes the action real and therefore explicitly removes the sign problem. By making comparisons with both the dual theory and also simulations at zero and imaginary chemical potential, the validity of the complex Langevin results are found to be correlated with the phase of the theory. Derived from these results, a set of tests are developed which are general and can be used to indicate whether the complex Langevin process has converged correctly.

Motivated by the incorrect convergence in the disordered part of the phase diagram, in Chapter 4 the XY model is rewritten in terms of components, rather than a spin variable. Although in principle both representations are equivalent, the complex Langevin processes converge differently leading to improved results with components.

A more serious and vexing problem with complex Langevin dynamics which has been observed since the early studies of the 1980s is that in some cases the complex Langevin process will converge to a stable distribution, but on closer inspection observables are incorrect, as was found with the XY model. Indeed, even in simple models of a single degree of freedom, where the exact result is readily available, the complex Langevin process can give wrong results. In Chapter 5 we cover a formal justification for the method and identify some points at which the argument might fail. From these we derive a necessary condition for correct results which can be checked numerically.

A detailed study of the SU(3) spin model with a chemical potential follows in Chapter 6, where the ideas and methods developed in the proceeding chapters are applied. Although this model also has a reformulation without a sign problem, at present it has not been used to generate results with which to compare. Nevertheless, the checks and tests which have been developed for complex Langevin dynamics all consistently suggest that the results are correct.

Chapter 2

Adaptive stepsize

The majority of this chapter is based on [57], of which I am the author. Where text is included for which I was not the original author, it is marked with a footnote.

2.1 Introduction

As has been well known since the 1980s, there are a number of problems associated with complex Langevin dynamics, see e.g. Refs. [44,46,47]. These can roughly be divided under two headings: instabilities and convergence. The first problem concerns the appearance of instabilities when solving the discretized Langevin equations numerically. Sometimes, but not always, this can be controlled by choosing a small enough stepsize. The second problem pertains to convergence. In some cases complex Langevin simulations appear to converge but not to the correct answer (see e.g. Ref. [47]). In order to disentangle these issues, we tackle in this chapter the first one and present adaptive stepsize algorithms that lead to a stable evolution and are not constrained to very small stepsizes only. A discussion of the second problem is deferred to subsequent chapters.

The chapter is organized as follows. In Section 2.2 we introduce the

problem, indicate why an adaptive stepsize might be necessary and outline the basic idea behind the algorithms, extending the ideas of Refs. [46,47]. To avoid notational cluttering we use a real scalar field, but we emphasize that the method is more generally applicable. We then present two algorithms implementing the basic idea, and apply them to the three-dimensional XY model at finite chemical potential in Section 2.3 and the heavy dense limit of QCD in Section 2.4. The latter was previously studied in [54]. We show a few selected results to indicate the applicability of the approach. The issue of convergence in the XY model is discussed in detail in Chapter 3 and in more generality in Chapter 5.

2.2 Adaptive stepsize

Consider a real scalar field ϕ with the Langevin equation of motion

$$\frac{\partial \phi}{\partial \vartheta} = -\frac{\delta S}{\delta \phi} + \eta. \tag{2.1}$$

Here ϑ is the supplementary Langevin time, $-\delta S/\delta\phi$ is the drift term derived from the action S, and η is Gaussian noise. The fundamental assertion of stochastic quantization is that in the infinite (Langevin) time limit, noise averages of observables become equal to quantum expectation values, defined via the standard functional integral,

$$\lim_{\vartheta\to\infty} \langle O[\phi(\vartheta)] \rangle_\eta = \frac{\int D\phi\, e^{-S} O[\phi]}{\int D\phi\, e^{-S}}, \tag{2.2}$$

where the brackets on the left denote a noise average.

If the action is complex the drift term becomes complex and so the field acquires an imaginary part (even if initially real). One must therefore com-

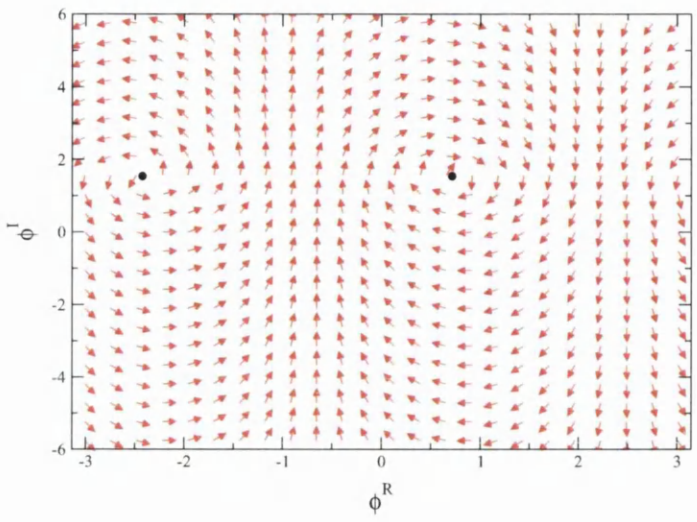


Figure 2.1: Example of a classical flow diagram in the XY model at nonzero chemical potential ($\mu = 2$). The arrows denote the normalized drift terms (K^R, K^I) at (ϕ^R, ϕ^I). The dots are classical fixed points.

plexify all fields, $\phi \to \phi^R + i\phi^I$. The Langevin equation then becomes

$$\frac{\partial \phi^R}{\partial \vartheta} = K^R + \eta, \qquad K^R = -\mathrm{Re}\,\frac{\delta S}{\delta \phi}\Big|_{\phi \to \phi^R + i\phi^I}, \tag{2.3}$$

$$\frac{\partial \phi^I}{\partial \vartheta} = K^I, \qquad K^I = -\mathrm{Im}\,\frac{\delta S}{\delta \phi}\Big|_{\phi \to \phi^R + i\phi^I}. \tag{2.4}$$

Here we restrict ourselves to real noise.

The complexification changes the dynamics substantially. Suppose that before complexification ϕ is a variable with a compact domain, e.g. $-\pi < \phi \leq \pi$. After complexification, the domain is noncompact since $-\infty < \phi^I < \infty$. Moreover, there will be unstable directions along which $\phi^I \to \pm\infty$. This is best seen in classical flow diagrams, in which the drift terms are plotted as a function of the degrees of freedom ϕ^R, ϕ^I. In Fig. 2.1 we show an example of a classical flow diagram in the XY model at finite chemical potential (to be discussed below). More examples of classical flow diagrams

with unstable directions can be found in Refs. [47, 53, 54, 59]. The arrows denote the drift terms $(K^{\mathrm{R}}, K^{\mathrm{I}})$ at $(\phi^{\mathrm{R}}, \phi^{\mathrm{I}})$. The length of the arrows is normalized for clarity. In this case there are unstable directions at $\phi^{\mathrm{R}} \sim -0.7$ and $\phi^{\mathrm{R}} \sim 2.4$. The black dots denote classical fixed points where the drift terms vanish. Generally speaking, the forces are larger when one is further away from the fixed points. In absence of the noise, one finds generically that configurations reach infinity in a finite time, since the forces grow exponentially for large imaginary field values.

When a Langevin trajectory makes a large excursion into imaginary directions, for instance by coming close to an unstable direction, sufficient care in the numerical integration of the Langevin equations is mandatory. In some cases it suffices to employ a small stepsize ϵ, after discretizing Langevin time as $\vartheta = n\epsilon$. However, this does not solve instabilities in all situations. Moreover, a small stepsize will result in a slow evolution, requiring many updates to explore configuration space.

To cure both problems, we introduce an adaptive stepsize, ϵ_n, in the discretized Langevin equations,

$$\phi_x^{\mathrm{R}}(n+1) = \phi_x^{\mathrm{R}}(n) + \epsilon_n K_x^{\mathrm{R}}(n) + \sqrt{\epsilon_n}\eta_x(n), \tag{2.5}$$

$$\phi_x^{\mathrm{I}}(n+1) = \phi_x^{\mathrm{I}}(n) + \epsilon_n K_x^{\mathrm{I}}(n), \tag{2.6}$$

where the noise satisfies

$$\langle \eta_x(n) \rangle = 0, \qquad \langle \eta_x(n)\eta_{x'}(n') \rangle = 2\delta_{xx'}\delta_{nn'}. \tag{2.7}$$

The magnitude of the stepsize is determined by controlling the distance a single update makes in the configuration space.

At each discrete Langevin time n, we monitor the quantity

$$K_n^{\max} = \max_x |K_x(n)| = \max_x \sqrt{K_x^{\mathrm{R}2}(n) + K_x^{\mathrm{I}2}(n)}. \tag{2.8}$$

We then place an upper bound on the product $\epsilon K^{\max}$ and define the stepsize ϵ_n as

$$\epsilon_n = \bar{\epsilon}\frac{\langle K^{\max}\rangle}{K_n^{\max}}. \tag{2.9}$$

Here $\bar{\epsilon}$ is the desired average stepsize (which can be controlled) and the expectation value of the maximum drift term $\langle K^{\max}\rangle$ is either precomputed, or computed during the thermalisation phase (with an initial guess). In this way, the stepsize is completely local in Langevin time and becomes smaller when the drift term is large (e.g. close to an instability) and larger when it is safe to do so.

In the second formulation[1], we keep $\epsilon K^{\max}$ within a factor p relative to a reference value $\mathcal{K}$, i.e.,

$$\frac{1}{p}\mathcal{K} \leq \epsilon K^{\max} \leq p\mathcal{K}. \tag{2.10}$$

If this range is exceeded the stepsize is increased/reduced by the factor p. This is iterated several times, if necessary. Both p and $\mathcal{K}$ have to be chosen beforehand, but this does not require fine tuning as long as clearly inadequate regions are avoided.

At this stage it is useful to identify the relationship between the adaptive stepsize procedure described above and the introduction of a kernel. With a field dependent kernel $\Gamma(x)$, the Langevin equation is modified to

$$dx = (\Gamma K + \Gamma')dt + \sqrt{\Gamma}dw, \tag{2.11}$$

with the noise term $dw = \sqrt{dt}\,\eta$. The extra term, Γ', is required in order for the Fokker-Planck equation to have the same fixed point as without a kernel. The derivation follows from the definition of expectation values,

$$\langle f(x)\rangle = \int f(x)P(x;t)dx. \tag{2.12}$$

[1]The second adaptive stepsize algorithm was developed and implemented by Stamatescu

Differentiating and Taylor expanding to first order in dt gives

$$\frac{\partial}{\partial t}\langle f(x)\rangle = \frac{1}{dt}\left[f(x) + f'(x)(\Gamma K + \Gamma')dt + \frac{1}{2}f''(x)\Gamma\,\eta^2\,dt - f(x)\right]. \quad (2.13)$$

Making use of the noise correlations and the definition of expectation values from (2.12) it follows that

$$\int \left[f'(x)(\Gamma K + \Gamma') + f''(x)\Gamma\right] P(x;t)dx = \int f(x)\frac{\partial}{\partial t}P(x;t). \quad (2.14)$$

After integrating by parts and rearranging terms it follows that

$$\frac{\partial}{\partial x}\Gamma\left(\frac{\partial}{\partial x} - K\right)P(x;t) = \frac{\partial}{\partial t}P(x;t), \quad (2.15)$$

which is the Fokker-Planck equation. Importantly, the inclusion of the extra kernel term Γ' does not alter the fixed point structure of the Fokker-Planck equation, which therefore still has the correct stationary solution of $P \sim e^{-S}$.

The adaptive stepsize algorithm described above can instead be cast as a kernel of the form

$$\Gamma(x) = \min\left\{1, \frac{\langle K^{\max}\rangle}{|K^{\max}|}\right\}. \quad (2.16)$$

However, the important difference between a correct implementation of a kernel and the adaptive stepsize procedure is the missing extra term in the Langevin equation, Γ'. Note that older literature which discusses the use of reducing the stepsize to stabilize the dynamics also neglects this term.

In a field theory, the missing term in the Langevin equation appears only on the site of $K^{\max}$ and its nearest neighbours. This will therefore be expected to have an effect of the order $O(1/\Omega)$ which will vanish in the thermodynamic limit.

At this stage it is important to note that computing noise averages is complicated by the inclusion of an adaptive stepsize. Some configurations will appear more frequently after each Langevin update due to the stepsize being

reduced to accommodate a large drift term. This means that if samples are taken after each timestep, the correct noise average when using an adaptive stepsize procedure is

$$\langle O \rangle = \frac{\sum_n O_n \epsilon_n}{\sum_n \epsilon_n}. \tag{2.17}$$

This procedure of reweighting observables with the timestep does not provide a fix for the lack of the kernel term in the Langevin equation.

Alternatively, one can avoid the issue of oversampling regions where the stepsize is small, taking samples after some fixed period of Langevin time instead of after each Langevin update. Provided this duration of time is sufficiently long to eliminate correlations between samples, this method should produce correct expectation values with respect to the configurations generated by the Langevin dynamics.

In the next sections, we apply these formulations to the XY model at nonzero chemical potential and QCD in the heavy dense limit respectively.

2.3 XY model

We demonstrate the first implementation using the three-dimensional XY model at finite chemical potential [30]. The action is

$$S = -\beta \sum_x \sum_{\nu=0}^{2} \cos\left(\phi_x - \phi_{x+\hat{\nu}} - i\mu\delta_{\nu,0}\right). \tag{2.1}$$

The theory is defined on a lattice of size $\Omega = N_\tau N_s^2$, with periodic boundary conditions in all three directions. The chemical potential is introduced as an imaginary constant vector field in the temporal direction [62] and couples to the conserved Noether charge associated with the global symmetry $\phi_x \to \phi_x + \alpha$. As always, the action is complex when $\mu \neq 0$ and satisfies $S^*(\mu) =$

$S(-\mu^*)$.[2]

The drift terms, after complexification, read

$$K_x^{\mathrm{R}} = -\beta \sum_\nu \Big[\sin\left(\phi_x^{\mathrm{R}} - \phi_{x+\hat{\nu}}^{\mathrm{R}}\right) \cosh\left(\phi_x^{\mathrm{I}} - \phi_{x+\hat{\nu}}^{\mathrm{I}} - \mu\delta_{\nu,0}\right) + \sin\left(\phi_x^{\mathrm{R}} - \phi_{x-\hat{\nu}}^{\mathrm{R}}\right) \cosh\left(\phi_x^{\mathrm{I}} - \phi_{x-\hat{\nu}}^{\mathrm{I}} + \mu\delta_{\nu,0}\right) \Big], \tag{2.2}$$

$$K_x^{\mathrm{I}} = -\beta \sum_\nu \Big[\cos\left(\phi_x^{\mathrm{R}} - \phi_{x+\hat{\nu}}^{\mathrm{R}}\right) \sinh\left(\phi_x^{\mathrm{I}} - \phi_{x+\hat{\nu}}^{\mathrm{I}} - \mu\delta_{\nu,0}\right) + \cos\left(\phi_x^{\mathrm{R}} - \phi_{x-\hat{\nu}}^{\mathrm{R}}\right) \sinh\left(\phi_x^{\mathrm{I}} - \phi_{x-\hat{\nu}}^{\mathrm{I}} + \mu\delta_{\nu,0}\right) \Big]. \tag{2.3}$$

As anticipated, they are unbounded due to the ϕ^{I} variables.

To construct the flow diagram in Fig. 2.1, we have chosen the field variables at the six sites neighbouring ϕ_x as random variables between $\pm\pi$. Note that the drift terms change sign when $\phi_x \to \phi_x + \pi$, for given x, explaining the symmetry in Fig. 2.1. The normalized drift terms and the classical fixed points ($K_x^{\mathrm{R}} = K_x^{\mathrm{I}} = 0$) are independent of β.

In an attempt to solve these Langevin equations numerically with a fixed stepsize, we found that instabilities and runaway trajectories appear so frequent, that it is practically impossible to construct a thermalized configuration, even when the stepsize is very small, say, $\epsilon \sim 10^{-5}$. This becomes worse on larger volumes.[3] We therefore switch to the adaptive scheme, using the first implementation. In Fig. 2.2 we show examples of the maximal drift term $K^{\max}$ and the adaptive stepsize ϵ_n as a function of Langevin time for three different lattice volumes and two values of β and μ. We observe that the maximal force fluctuates over several orders of magnitude during the evo-

[2]At nonzero chemical potential, it is preferable to interpret this system as a three-dimensional Euclidean quantum field theory at finite temperature (with coupling β and inverse temperature N_τ), rather than as a three-dimensional classical spin system with inverse temperature β. Models in this class can also be studied using world-line formulations [29,30].

[3]This is in sharp contrast with the relativistic Bose gas in Ref. [55], where instabilities were not encountered for the parameter values used there.

lution (note the vertical logarithmic scale). Moreover, the frequency and size of the fluctuations increase when increasing the lattice volume. This is consistent with the picture developed above: on a larger volume there are more opportunities to be on a potentially unstable trajectory and subject to large forces. The effect also gets worse at larger chemical potential. $K^{\max}$ and ϵ_n are inversely proportional, as expected. We note that although it is necessary to use a tiny stepsize occasionally, the algorithm is designed such that the evolution will continue with a larger stepsize as soon as possible. As a result, the time average of ϵ_n is close to the input timestep $\bar{\epsilon} = 0.01$ in all cases.[4] We emphasize that after the implementation of this algorithm we have not observed any instabilities, for a wide range of parameters ($0.1 < \beta < 2$, $0 < \mu < 6$), lattice sizes (up to 16^3), and long runtimes (we explored millions of timesteps, corresponding to Langevin times of several thousand).

To illustrate the method, we introduce two related models with a real action: the XY model with imaginary chemical potential $\mu = i\mu_{\mathrm{I}}$, with the action

$$S_{\mathrm{imag}} = -\beta \sum_{x,\nu} \cos\left(\phi_x - \phi_{x+\hat{\nu}} + \mu_{\mathrm{I}}\delta_{\nu,0}\right), \tag{2.4}$$

and the phase quenched theory, obtained by taking the absolute value of the complex weight, i.e. $e^{-S} \rightarrow |e^{-S}| \equiv e^{-S_{\mathrm{pq}}}$, which yields the action

$$S_{\mathrm{pq}} = -\beta \sum_{x,\nu} \cos\left(\phi_x - \phi_{x+\hat{\nu}}\right) \cosh\left(\mu\delta_{\nu,0}\right). \tag{2.5}$$

This is the anisotropic XY model, with direction-dependent coupling β_ν, where $\beta_0 = \beta\cosh\mu$ and $\beta_i = \beta$ $(i = 1, 2)$. Both models are solved using real Langevin dynamics. The drift terms are bounded and there are no instabilities.

In Fig. 2.3 we show the expectation value of the action density $\langle S\rangle/\Omega$ in the high-β phase, at $\beta = 0.55$, for small values of μ^2.[5] The lattice sizes

[4] In practice, one may take $\epsilon_n \leq \bar{\epsilon}$ always, to prevent the appearance of large timesteps.

[5] Recall that at zero chemical potential, the three-dimensional XY model has a contin-

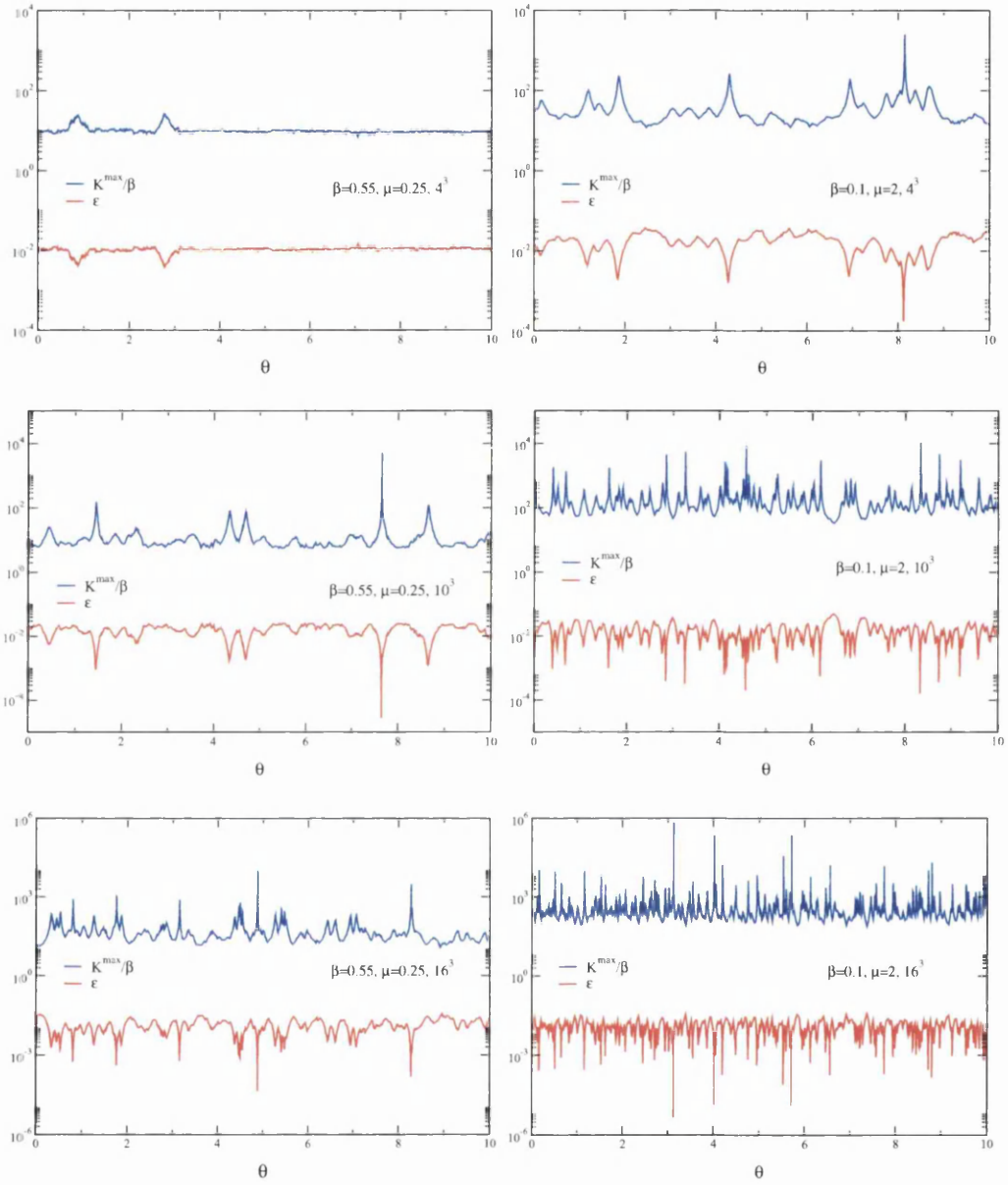


Figure 2.2: Example of the Langevin evolution of the maximal drift term $K^{\max}/\beta$ and the adaptive stepsize ϵ_n in the three-dimensional XY model with $\beta = 0.55$ and $\mu = 0.25$ (left) and $\beta = 0.1$ and $\mu = 2$ (right) on lattices of size 4^3 (top), 10^3 (middle) and 16^3 (bottom). The input stepsize is $\bar{\epsilon} = 0.01$. Note the vertical logarithmic scale.

are 6^3 and 8^3, showing that finite size effects are under control. The result at imaginary μ appears at $\mu^2 < 0$, while the complex and phase quenched results are plotted at $\mu^2 > 0$. At $\mu = 0$ all results agree (within the statistical error). The complex and phase quenched results differ, as can be expected from e.g. a Taylor expansion of the observable for small μ. The results for imaginary and real μ appear continuous around $\mu^2 = 0$, which is expected from the analyticity of the partition function in μ^2 on a finite lattice. The lines indicate second-order fits to the data on the 6^3 lattice, with the results

$$\langle S \rangle / \Omega = -0.9433(7) - 0.502(4)\mu^2 + 0.19(1)\mu^4, \tag{2.6}$$

$$\langle S \rangle_{\mathrm{pq}} / \Omega = -0.940(2) \quad - 0.35(2)\mu^2 \quad - 0.04(3)\mu^4. \tag{2.7}$$

In the first case, the data at real and imaginary μ are combined in the fit.

For imaginary μ we observe a cusp at $\mu_{\mathrm{I}} = \pi / N_\tau$. This is similar to the Roberge-Weiss transition in QCD [6, 64] and is due to the periodicity $\mu_{\mathrm{I}} \to \mu_{\mathrm{I}} + 2\pi / N_\tau$.[6] Note that the dashed lines in Fig. 2.3 reflect this symmetry. It would therefore be interesting to determine the phase structure of the XY model at imaginary chemical potential and finite N_τ. In the three-dimensional thermodynamic limit (N_τ is taken to infinity as well) and vanishing chemical potential, the magnetized and symmetric phase are separated at the critical coupling $\beta_c(\mu = 0)$. Consequently it is intriguing to analyse the interplay between the putative Roberge-Weiss transition and the standard magnetization transition in this limit, in particular since it would result in a breakdown of analyticity of $\beta_c(\mu^2)$ around $\mu^2 = 0$ and make $\beta_c(\mu = 0)$ a multicritical point.

---

uous phase transition at $\beta_c(\mu = 0) \approx 0.4542$ (see e.g. Ref. [63]), separating the symmetric phase at small β from the symmetry broken phase at large β.

[6] One way to see this is by using a field redefinition, $\phi_x = \phi'_x + \mu_{\mathrm{I}} \tau$, which moves the μ dependence to the boundary condition $\phi'_{N_\tau, \mathbf{x}} = \phi'_{0, \mathbf{x}} - \mu_{\mathrm{I}} N_\tau$, similar as in fermionic models. The centre symmetry is of course trivial in this model.

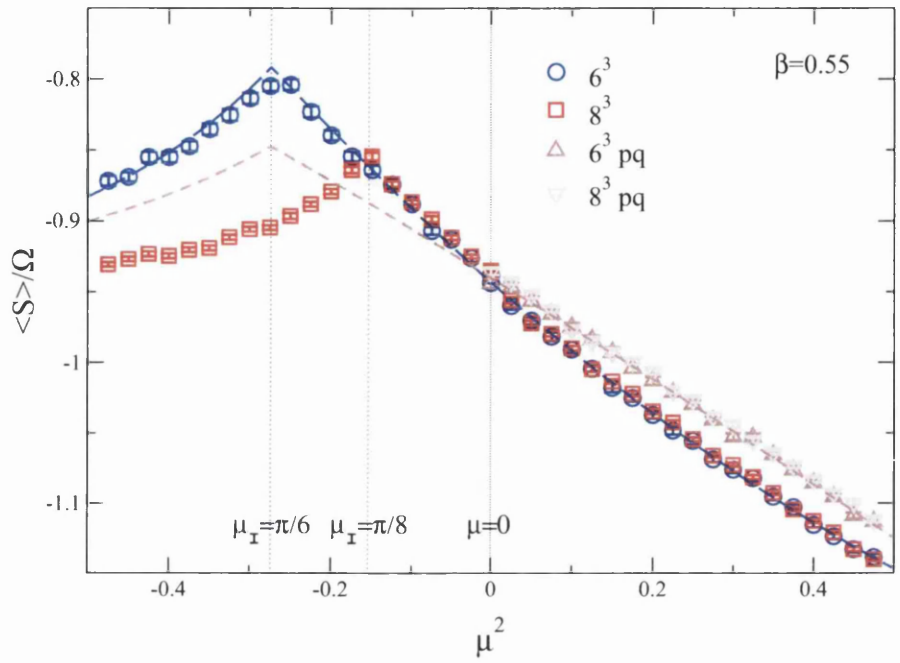


Figure 2.3: Action density $\langle S\rangle/\Omega$ in the three-dimensional XY model as a function of μ^2 at $\beta = 0.55$ on lattices of size 6^3 and 8^3, for the full theory (circle, square, $\mu^2 > 0$), at imaginary μ (circle, square, $\mu^2 < 0$), and phase quenched (triangles, $\mu^2 > 0$). The vertical lines at $\mu_{\rm I} = \pi/N_\tau$ indicate the Roberge-Weiss lines at imaginary μ. The dashed lines are the second-order fits (2.6, 2.7), incorporating the RW reflection symmetry.

2.4 Heavy dense limit of QCD

To show the generality of the adaptive stepsize method we now apply it to the heavy dense limit of QCD in four dimensions[7]. Here we shall present results obtained with the second algorithm. The stochastic quantization for this theory was studied in Ref. [54]. Here we briefly repeat the essential equations; we refer to Ref. [54] for further details.

[7]This section, which is the work of Stamatescu, formed a part of the original paper [57] and is included here for completeness.

The gluonic part of the action is the standard Wilson SU(3) action,

$$S_B[U] = -\beta \sum_x \sum_{\substack{\mu,\nu \\ \mu<\nu}} \left(\frac{1}{6} \left[\mathrm{Tr}\, U_{x,\mu\nu} + \mathrm{Tr}\, U_{x,\mu\nu}^{-1} \right] - 1 \right), \tag{2.1}$$

where $U_{x,\mu\nu}$ are the plaquettes and $\beta = 6/g^2$. The fermion determinant (starting from Wilson fermions) is approximated as

$$\det M \approx \prod_{\mathbf{x}} \det \left(1 + h e^{\mu/T} \mathcal{P}_{\mathbf{x}}\right)^2 \det \left(1 + h e^{-\mu/T} \mathcal{P}_{\mathbf{x}}^{-1}\right)^2, \tag{2.2}$$

where $h = (2\kappa)^{N_\tau}$. Here κ is the Wilson hopping parameter and $N_\tau = 1/T$ the number of sites in the temporal direction. The lattice spacing $a \equiv 1$. The determinant refers to colour space only. The (conjugate) Polyakov loops are

$$\mathcal{P}_{\mathbf{x}} = \prod_{\tau=0}^{N_\tau - 1} U_{(\tau,\mathbf{x}),4}, \qquad \mathcal{P}_{\mathbf{x}}^{-1} = \prod_{\tau=N_\tau-1}^{0} U_{(\tau,\mathbf{x}),4}^{-1}. \tag{2.3}$$

A formal derivation of Eq. (2.2) follows by considering the heavy ($\kappa \to 0$) and dense ($\mu \to \infty$) limit, keeping the product κe^μ fixed (see Ref. [66] and references therein). The anti-quark contribution is kept to preserve the symmetry under complex conjugation.

The Langevin process is

$$U'_{x,\mu} = R_{x,\mu}\, U_{x,\mu}, \qquad R_{x,\mu} = \exp\left[i\lambda_a \left(\epsilon K_{x\mu a} + \sqrt{\epsilon}\eta_{x\mu a} \right) \right], \tag{2.4}$$

where λ_a are the Gell-Mann matrices and the noise satisfies $\langle \eta_{x\mu a} \rangle = 0$, $\langle \eta_{x\mu a} \eta_{y\nu b} \rangle = 2\delta_{\mu\nu}\delta_{ab}\delta_{xy}$. The drift term $K_{x\mu a} = -D_{x\mu a}\,(S_B + S_F)$, where $S_F = -\ln \det M$, is complex due to the fermion contribution. For explicit expressions, see Ref. [54].

This Langevin process suffers from instabilities, which can be partially controlled using small enough stepsize. Applying the second formulation of

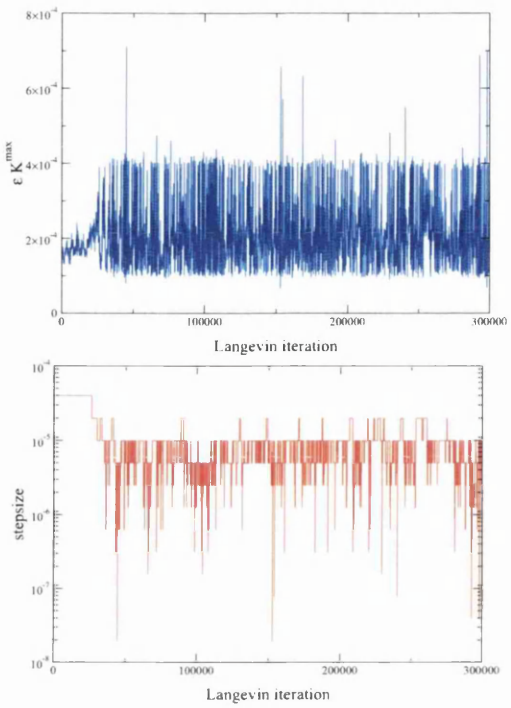


Figure 2.4: Example of the Langevin evolution of the maximal drift term $\epsilon K^{\max}$ (left) and the adaptive stepsize ϵ (right) in the heavy dense limit of QCD with $\beta = 5$, $\kappa = 0.12$ and $\mu = 0.7$ on a lattice of size 2^4, using $\mathcal{K} = 2 \times 10^{-4}$.

the adaptive stepsize algorithm completely eliminates runaways. In this setup measurements are performed after a Langevin time interval of length $\Delta\vartheta$ (this defines one "iteration"). The number of sweeps N_{sweep} in an iteration depends on the Langevin timestep ϵ: if the latter is decreased by a factor p, N_{sweep} is increased by the same factor, and conversely. This makes the statistical analysis straightforward, since each iteration has the same weight $\Delta\vartheta = \epsilon\, N_{\text{sweep}}$. Note that N_{sweep} will no longer be decreased and correspondingly ϵ will no longer be increased if the former reaches 1.

To demonstrate this approach, we show in Fig. 2.4 a characteristic evo-

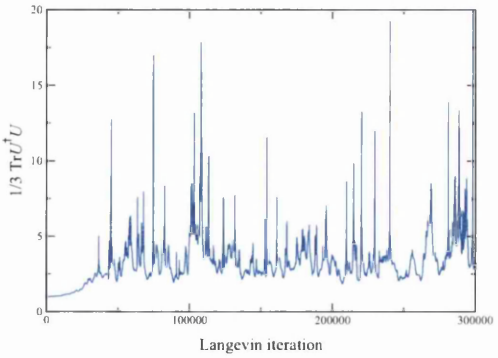


Figure 2.5: As above for $\frac{1}{3}\mathrm{Tr}\, U_4 U_4^\dagger$, indicating the deviation from unitarity during the evolution.

lution of $\epsilon K^{\max}$ (left) and the stepsize ϵ (right), using $\mathcal{K} = 2 \times 10^{-4}$ and $p = 2$, with initial $\epsilon = 1 \times 10^{-5}$, $N_{\mathrm{sweep}} = 16$. The total number of iterations is $N_{\mathrm{iter}} = 3 \times 10^5$ leading to a total Langevin time $\vartheta_{\mathrm{tot}} = N_{\mathrm{iter}} \times \Delta\vartheta = 48$. Varying these input parameters by factors of two or more does not change the results but may only affect the statistics. $K^{\max}$ is the maximum value of $|K_{x\mu a}|$ over x, μ and a in the last sweep of an iteration. We observe that the product $\epsilon K^{\max}$ remains bounded, as required, while the stepsize (and hence $K^{\max}$) fluctuate substantially. In Fig. 2.5 we show the evolution of $\frac{1}{3}\mathrm{Tr}\, U_4 U_4^\dagger$, which measures the deviation from unitarity [54], acknowledging the typical fluctuations in stationary regime.

As mentioned above, after the implementation of this algorithm, we have not encountered any instability.

2.5 Conclusion

Instabilities can make complex Langevin simulations extremely problematic. We have shown that they result from large fluctuations in the magnitude of the drift term and the presence of unstable directions in complexified field

space. This can be cured by using an adaptive time-local stepsize. The scheme is generic and could be applied to other theories such as QCD. While we have no proof that stable evolution is guaranteed, we have not encountered any instability using this method in the XY model at finite chemical potential and QCD in the heavy dense limit, for a wide selection of parameter values and system sizes. With a fixed stepsize on the other hand, instabilities can appear so frequently that it is virtually impossible to generate a thermalized ensemble.

This adaptive stepsize procedure can be written as a kernel. However, the addition of a kernel requires an extra term to be introduced into the Langevin equation in order to preserve the correct stationary solution. Although this term was neglected, in the case of a field theory its effect vanishes in the thermodynamic limit, since it appears only as $O(1/\Omega)$. However, it is less clear that in simple models of a single variable that the extra term can be ignored. Nevertheless, empirical results suggest that the algorithm works correctly.

Runaways are due to specific instabilities of the complexified Langevin equations caused by the strong increase of the drift in the non-compact imaginary directions. We have shown that these runaways can be eliminated by using a dynamical step size, which indicates that they are not a question of principle for complex Langevin dynamics but one of numerical accuracy in following the trajectories. We consider this to be an important result of our analysis. Of course, for practical calculations one should consider further optimization of the algorithms by applying methods developed for general real Langevin processes (see, e.g., Ref. [67]) after analysing their adequacy to the complex Langevin problems of interest. Note it is known that some cases, such as slave equations, lead to a distribution with a mean but no variance. The repulsive runaway trajectories are important and cannot be ignored.

We emphasize that the adaptive stepsize permits a fine tracing of the drift trajectories. If the process picks up a diverging trajectory the noise term will

typically kick the process off it. The present results suggest therefore that runaways are not due to following diverging trajectories [58] but rather due to following "wrong" trajectories, i.e. trajectories which, because of accumulating errors in the evaluation of the drift, do not belong to the dynamics of the problem. This both stresses the necessity of ensuring precision in the calculation and helps in disentangling various sources of deficiency in the search for a reliable method.

Chapter 3

XY model

3.1 Introduction

In this chapter, we continue our investigation into the applicability of complex Langevin dynamics at finite chemical potential [54–58]. We consider the three-dimensional XY model for a number of reasons. In Chapter 2 it was shown that this theory is very sensitive to instabilities and runaways and therefore requires the use of an adaptive stepsize [57]. This is similar to the case of QCD in the heavy dense limit [54, 57]. As with QCD, this theory has a Roberge-Weiss periodicity at imaginary chemical potential [57, 64]. Furthermore, as shown in Chapter 4, the XY model is closely related to the relativistic Bose gas at finite chemical potential, for which complex Langevin dynamics was shown to work very well (at weak coupling in four dimensions) [55, 56]. Finally, this theory can be rewritten using a world line formulation without a sign problem [30, 31], which can be solved efficiently using the worm algorithm [31,68]. This allows for a direct comparison for all parameter values.

The chapter is organized as follows. In Sec. 3.2, we remind the reader of some details of the XY model at real and imaginary chemical potential, the adaptive stepsize algorithm we use and the related phase-quenched XY

model. The world line formulation and some properties of the strong-coupling expansion are briefly mentioned in Sec. 3.3. We then test the validity of complex Langevin dynamics in Sec. 3.4 and develop diagnostic tests in Sec. 3.5.

3.2 XY model

The action of the XY model at finite chemical potential is

$$S = -\beta \sum_x \sum_{\nu=0}^{2} \cos(\phi_x - \phi_{x+\hat{\nu}} - i\mu\delta_{\nu,0}), \tag{3.1}$$

where $0 \leq \phi_x < 2\pi$. The theory is defined on a lattice of size $\Omega = N_\tau N_s^2$, and we use periodic boundary conditions. The chemical potential μ is coupled to the Noether charge associated with the global symmetry $\phi_x \to \phi_x + \alpha$ and is introduced in the standard way [62]. The action satisfies $S^*(\mu) = S(-\mu^*)$. At vanishing chemical potential the theory is known to undergo a phase transition at $\beta_c = 0.45421$ [31, 63] between a disordered phase when $\beta < \beta_c$ and an ordered phase when $\beta > \beta_c$.

The drift terms appearing in the complex Langevin equations are given by

$$\begin{aligned} K_x^{\mathrm{R}} = -\beta \sum_\nu \Big[& \sin(\phi_x^{\mathrm{R}} - \phi_{x+\hat{\nu}}^{\mathrm{R}}) \cosh(\phi_x^{\mathrm{I}} - \phi_{x+\hat{\nu}}^{\mathrm{I}} - \mu\delta_{\nu,0}) \\ & + \sin(\phi_x^{\mathrm{R}} - \phi_{x-\hat{\nu}}^{\mathrm{R}}) \cosh(\phi_x^{\mathrm{I}} - \phi_{x-\hat{\nu}}^{\mathrm{I}} + \mu\delta_{\nu,0}) \Big], \end{aligned} \tag{3.2a}$$

$$\begin{aligned} K_x^{\mathrm{I}} = -\beta \sum_\nu \Big[& \cos(\phi_x^{\mathrm{R}} - \phi_{x+\hat{\nu}}^{\mathrm{R}}) \sinh(\phi_x^{\mathrm{I}} - \phi_{x+\hat{\nu}}^{\mathrm{I}} - \mu\delta_{\nu,0}) \\ & + \cos(\phi_x^{\mathrm{R}} - \phi_{x-\hat{\nu}}^{\mathrm{R}}) \sinh(\phi_x^{\mathrm{I}} - \phi_{x-\hat{\nu}}^{\mathrm{I}} + \mu\delta_{\nu,0}) \Big]. \end{aligned} \tag{3.2b}$$

The equations are integrated numerically by discretizing Langevin time as

$\vartheta = n\epsilon_n$ with ϵ_n the adaptive stepsize. Explicitly,

$$\phi_x^{\mathrm{R}}(n+1) = \phi_x^{\mathrm{R}}(n) + \epsilon_n K_x^{\mathrm{R}}(n) + \sqrt{\epsilon_n}\eta_x(n), \tag{3.3a}$$

$$\phi_x^{\mathrm{I}}(n+1) = \phi_x^{\mathrm{I}}(n) + \epsilon_n K_x^{\mathrm{I}}(n), \tag{3.3b}$$

where we specialized to real noise, with $\langle \eta_x(n)\eta_{x'}(n')\rangle = 2\delta_{xx'}\delta_{nn'}$. In the case that $\mu = \phi^{\mathrm{I}} = 0$, the drift terms are bounded and $|K_x^{\mathrm{R}}| < 6\beta$. When $\phi^{\mathrm{I}} \neq 0$, the drift terms are unbounded, which can result in instabilities and runaways. In this particular theory, much care is required to numerically integrate the dynamics in a stable manner and we found that an adaptive stepsize is mandatory [57]. At each timestep, the stepsize is determined according to

$$\epsilon_n = \min\left\{\bar{\epsilon}, \bar{\epsilon}\frac{\langle K^{\max}\rangle}{K_n^{\max}}\right\}, \tag{3.4}$$

where

$$K_n^{\max} = \max_x \left|K_x^{\mathrm{R}}(n) + iK_x^{\mathrm{I}}(n)\right|. \tag{3.5}$$

Here $\bar{\epsilon}$ is the desired target stepsize and $\langle K^{\max}\rangle$ is either precomputed or computed during the thermalisation phase. All observables are analyzed over equal periods of Langevin time to ensure correct statistical significance, as discussed in Chapter 2.

The observable we focus on primarily in this study is the action density $\langle S\rangle/\Omega$. After complexification the action is written as $S = S^{\mathrm{R}} + iS^{\mathrm{I}}$, with

$$S^{\mathrm{R}} = -\beta\sum_{x,\nu}\cos(\phi_x^{\mathrm{R}} - \phi_{x+\hat{\nu}}^{\mathrm{R}})\cosh(\phi_x^{\mathrm{I}} - \phi_{x+\hat{\nu}}^{\mathrm{I}} - \mu\delta_{\nu,0}), \tag{3.6a}$$

$$S^{\mathrm{I}} = \beta\sum_{x,\nu}\sin(\phi_x^{\mathrm{R}} - \phi_{x+\hat{\nu}}^{\mathrm{R}})\sinh(\phi_x^{\mathrm{I}} - \phi_{x+\hat{\nu}}^{\mathrm{I}} - \mu\delta_{\nu,0}). \tag{3.6b}$$

After noise averaging, the expectation value of the imaginary part is consistent with zero while the expectation value of the real part is even in μ, as is expected from symmetry considerations.

By choosing an imaginary chemical potential $\mu = i\mu_{\rm I}$ the action (3.1) becomes purely real. This has both the advantage of enabling standard Monte Carlo algorithms to be applied (we choose to employ real Langevin dynamics) and that the behaviour at $\mu^2 \gtrsim 0$ can be assessed by continuation from $\mu^2 \lesssim 0$. The action and drift term with imaginary chemical potential are

$$S_{\rm imag} = -\beta \sum_{x,\nu} \cos(\phi_x - \phi_{x+\hat{\nu}} + \mu_{\rm I}\delta_{\nu,0}), \tag{3.7}$$

$$K_x = -\beta \sum_{\nu} \left[\sin(\phi_x - \phi_{x+\hat{\nu}} + \mu_{\rm I}\delta_{\nu,0}) + \sin(\phi_x - \phi_{x-\hat{\nu}} - \mu_{\rm I}\delta_{\nu,0})\right]. \tag{3.8}$$

This theory is periodic under $\mu_{\rm I} \rightarrow \mu_{\rm I} + 2\pi/N_\tau$, which yields a Roberge-Weiss transition at $\mu_{\rm I} = \pi/N_\tau$, similar to QCD [64]. This periodicity can be made explicit by shifting the chemical potential to the final time slice, via the field redefinition $\phi_{\mathbf{x},\tau} \longrightarrow \phi'_{\mathbf{x},\tau} = \phi_{\mathbf{x},\tau} - \mu_{\rm I}\tau$. The action is then (for arbitrary complex chemical potential)

$$S_{\rm fts} = -\beta \sum_{x,\nu} \cos(\phi_x - \phi_{x+\hat{\nu}} - iN_\tau\mu\delta_{\tau,N_\tau}\delta_{\nu,0}). \tag{3.9}$$

We have also carried out simulations with this action and confirmed the results obtained with the original formulation. The sole exception was the largest β value ($\beta = 0.7$), where the original action missed the Roberge-Weiss transition, while the final-time-slice formulation located it without problems.

The severity of the sign problem is conventionally (see e.g. Ref. [1]) estimated by the expectation value of the phase factor $e^{i\varphi} = e^{-S}/|e^{-S}|$ in the phase quenched theory, i.e. in the theory where only the real part of the action (3.1) is included in the Boltzmann weight. In this case, the phase quenched theory is the anisotropic XY model, with the action

$$S_{\rm pq} = -\sum_{x,\nu} \beta_\nu \cos(\phi_x - \phi_{x+\hat{\nu}}), \tag{3.10}$$

where $\beta_0 = \beta \cosh \mu$, and $\beta_{1,2} = \beta$.

3.3 World line formulation

The advantage of the XY model is that it can be formulated without a sign problem by an exact rewriting of the partition function in terms of world lines [30, 31].[1] Moreover, this dual formulation can be simulated efficiently with a worm algorithm [31, 68], which allows us to compare the results obtained with complex Langevin dynamics with those from the world line approach. We briefly repeat some essential elements of the world line formulation and refer to Ref. [31] for more details. The partition function can be rewritten using the identity

$$e^{\beta \cos \phi} = \sum_{k=-\infty}^{\infty} I_k(\beta) e^{ik\phi}, \tag{3.1}$$

where $I_k(\beta)$ are the modified Bessel functions of the first kind. Using this replacement and integrating over the fields, the partition function is written as

$$Z = \int D\phi \, e^{-S} = \sum_{[k]} \prod_{x,\nu} I_{k_{x,\nu}}(\beta) e^{k_{x,\nu} \mu \delta_{\nu,0}} \delta \left(\sum_{\nu} [k_{x,\nu} - k_{x-\hat{\nu},\nu}] \right). \tag{3.2}$$

The sum over $[k]$ indicates a sum over all possible world line configurations. Since $\langle S \rangle = -\beta \frac{\partial \ln Z}{\partial \beta}$, the action can be computed from

$$\langle S \rangle = -\beta \left\langle \sum_{x,\nu} \left[\frac{I_{k_{x,\nu}-1}(\beta)}{I_{k_{x,\nu}}(\beta)} - \frac{k_{x,\nu}}{\beta} \right] \right\rangle_{\text{wl}}, \tag{3.3}$$

where the brackets denote the average over world line configurations. To compute this average, we have implemented the worm algorithm, following

[1] The world line formulation has of course a long history in lattice gauge theory, see e.g. Ref. [69]. Recent work includes Refs. [70–72]. For a review, see Ref. [30].

Ref. [31]. We note here, amusingly, that the world line formulation has a sign problem at imaginary chemical potential.

Inspired by Ref. [73], we have also studied a (low-order) strong-coupling expansion of this model, using

$$I_k(2x) = \frac{x^k}{k!}\left(1 + \frac{x^2}{k+1} + \frac{x^4}{2(k+2)(k+1)} + \ldots\right). \tag{3.4}$$

At strong coupling the chemical potential cancels in most world lines, except when the world line wraps around the temporal direction. At leading order in the strong-coupling expansion, it then appears in the combination $(\frac{1}{2}\beta e^{\mu})^{N_\tau}$. In the thermodynamic limit it therefore contributes only when $\frac{1}{2}\beta e^{\mu} \geq 1$. Hence a simple strong-coupling estimate for the critical coupling at nonzero μ is given by

$$\beta_c(\mu) = 2e^{-\mu}. \tag{3.5}$$

The μ-independence at small β and μ is known as the Silver Blaze feature in QCD [74].

The partition function is expressed in terms of the free energy density f as $Z = \exp(-\Omega f)$. A strong-coupling expansion to order β^4 on a lattice with $N_\tau > 4$ yields

$$f = -\frac{3}{4}\beta^2 - \frac{21}{64}\beta^4 + \mathcal{O}(\beta^6), \tag{3.6}$$

and hence

$$\langle S\rangle/\Omega = -\frac{3}{2}\beta^2 - \frac{21}{16}\beta^4 + \mathcal{O}(\beta^6). \tag{3.7}$$

In the phase quenched theory we find

$$f_{\rm pq} = -\frac{1}{4}\beta^2\left(2 + \cosh^2\mu\right) - \frac{1}{64}\beta^4\left(14 + 8\cosh^2\mu - \cosh^4\mu\right) + \mathcal{O}(\beta^6). \tag{3.8}$$

We can now estimate the severeness of the sign problem at strong coupling.

The average phase factor takes the standard form,

$$\langle e^{i\varphi} \rangle_{\rm pq} = \frac{Z}{Z_{\rm pq}} = \exp\left[-\Omega \Delta f\right], \qquad \Delta f = f - f_{\rm pq}, \tag{3.9}$$

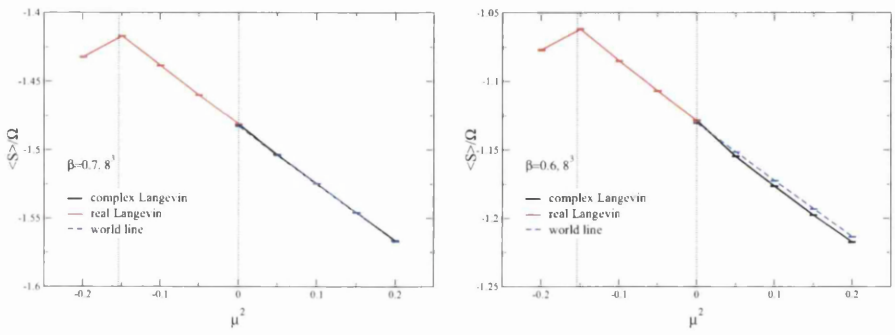
where in this case

$$\Delta f = \frac{1}{4}\beta^2 \left(\cosh^2 \mu - 1\right) + \frac{1}{64}\beta^4 \left(\cosh^2 \mu - 1\right)\left(7 - \cosh^2 \mu\right) + \mathcal{O}(\beta^6). \tag{3.10}$$

On a finite lattice and for small chemical potential we find therefore the sign problem to be mild in the strong-coupling limit, since the volume factor is balanced by $\beta^2\mu^2/4 \ll 1$.

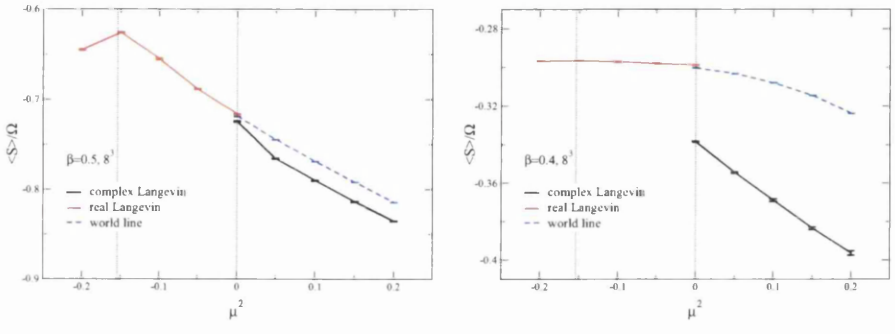
3.4 Comparison

We start to assess the applicability of complex Langevin dynamics for this model at small chemical potential. In this case we can use continuity arguments to compare observables at real and imaginary chemical potential. In Fig. 3.1 the real part of the action density is shown as a function of μ^2, for several values of β: from the ordered phase at large β to the disordered phase at low β. We observe that at the highest values of β this observable is continuous across $\mu^2 = 0$, which is a good indication that complex Langevin dynamics works well in this region. The cusp at $\mu_{\rm I} = \pi/N_\tau$ (corresponding to $\mu^2 = -0.154$) reflects the Roberge-Weiss transition. At lower β, however, we observe that the action density is no longer continuous: this is interpreted as a breakdown of complex Langevin dynamics. In order to verify this, Fig. 3.1 also contains the expectation values of the action density found using the worm algorithm in the world line formalism for real μ. As expected, in this case the action density is continuous across $\mu^2 = 0$ for all values of β, confirming the interpretation given above. We have verified that the jump in the action density at lower β is independent of the lattice volume. We have also verified that the discrepancy at $\mu^2 = 0$ between real Langevin dynamics and



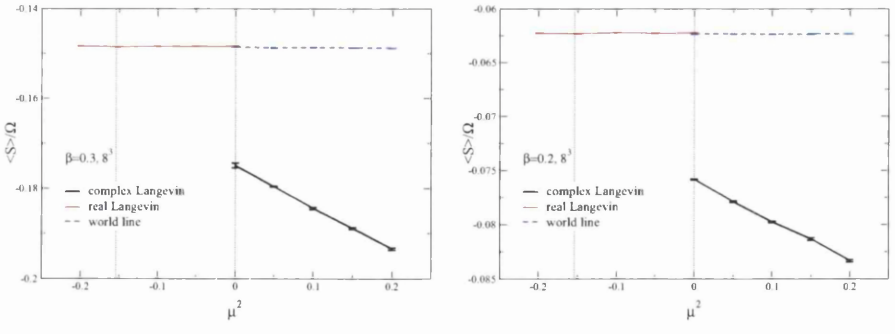
(a) $\beta = 0.7$

(b) $\beta = 0.6$



(c) $\beta = 0.5$

(d) $\beta = 0.4$



(e) $\beta = 0.3$

(f) $\beta = 0.2$

Figure 3.1: Real part of action density $\langle S \rangle / \Omega$ as a function of μ^2 on a lattice of size 8^3, using complex Langevin dynamics and the world line formulation at real μ ($\mu^2 > 0$) and real Langevin dynamics at imaginary μ ($\mu^2 < 0$). The vertical lines on the left indicate the Roberge-Weiss transitions at $\mu_{\mathrm{I}} = \pi/8$.

the world line result (at e.g. $\beta = 0.4$) is due to the finite Langevin stepsize.

For small β, the numerical results found with the worm algorithm are consistent with those derived analytically in the strong-coupling limit above. The expectation value of the action density is μ-independent and hence the Roberge-Weiss periodicity is smoothly realized. Using Eq. (3.7), we also find quantitative agreement: in the strong-coupling expansion $\langle S \rangle / \Omega = -0.0621 + \mathcal{O}(10^{-4})$ for $\beta = 0.2$ and $-0.145 + \mathcal{O}(10^{-3})$ for $\beta = 0.3$.

As discussed above, for the parameter values and lattice sizes used here the sign problem is not severe: taking $\mu^2 = 0.1$ and $\beta = 0.2$, we find that

$$\Omega \Delta f \approx \Omega \frac{\beta^2 \mu^2}{4} \approx 0.51, \qquad \langle e^{i\varphi} \rangle_{\mathrm{pq}} \approx 0.60. \tag{3.1}$$

We take this as a first indication that the observed breakdown is not due to the presence of the sign problem, especially since complex Langevin dynamics has been demonstrated to work well in other models where the sign problem is severe [55, 56].

To probe the reliability of complex Langevin dynamics for larger values of μ, we have computed the action density for a large number of parameter values in the $\beta - \mu$ plane. Our findings are summarized in Fig. 3.2, where we show the relative difference between the action densities obtained with complex Langevin (cl) and in the world line formulation (wl), according to

$$\Delta S = \frac{\langle S \rangle_{\mathrm{wl}} - \langle S \rangle_{\mathrm{cl}}}{\langle S \rangle_{\mathrm{wl}}}. \tag{3.2}$$

Also shown in this figure is the phase transition line $\beta_c(\mu)$, taken from Ref. [31]. We observe a clear correlation between the breakdown of Langevin dynamics and the phase boundary: complex Langevin dynamics works fine well inside the ordered phase, but breaks down in the boundary region and the disordered phase. The largest deviation around $\mu = 2$ is due to the Silver Blaze effect: the difference between the action density found with complex

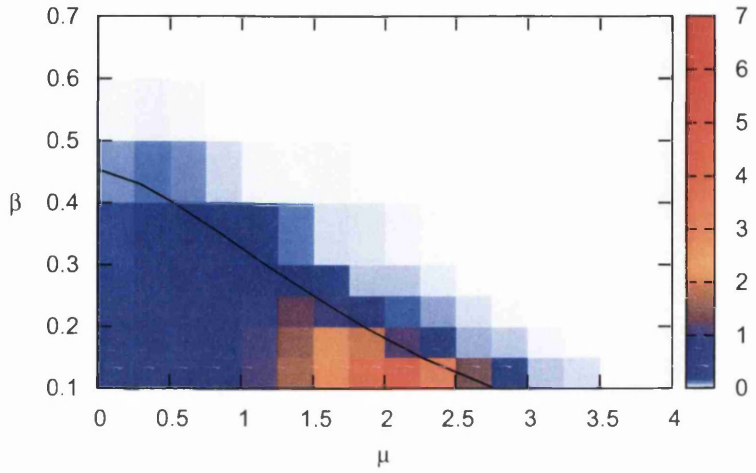


Figure 3.2: Colour plot indicating the relative difference ΔS between the expectation value of the action density obtained with complex Langevin dynamics and in the world line formulation, see Eq. (3.2). Also shown is the phase boundary $\beta_c(\mu)$ between the ordered (large β) and disordered (small β) phase [31].

Langevin dynamics and the correct μ-independent action density is maximal just before crossing over to the other phase, where the agreement improves quickly.

3.5 Diagnostics

In this section we attempt to characterize the results presented above in terms of properties of complex Langevin dynamics and the distribution $P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}]$ in the complexified field space, see Eq. (1.27). We suppress Langevin time dependence, since we always consider the quasi-stationary regime, i.e. the initial part of the evolution is discarded (we considered Langevin times up

to $\vartheta \sim 2 \times 10^4$). Our aim is to argue that the discrepancy at small β is introduced by complex Langevin dynamics rather than by the presence of a chemical potential and hence not due to the sign problem.

A first test of the validity of complex Langevin dynamics is to compare simulations at $\mu = 0$ using a cold start, i.e. with $\phi^{\mathrm{I}} = 0$ initially, and a hot start in which ϕ^{I} is taken from a Gaussian distribution.[2] When $\mu = 0$, a cold start corresponds to real Langevin dynamics. In the case of a hot start, however, the fields lie immediately in the complexified space and so the dynamics is complexified. Comparison of results obtained with these two initial ensembles gives insight into the inner workings of complex Langevin dynamics. We have computed the expectation value of the action density at $\mu = 0$ using both a hot and a cold start. We found them to agree at large β, despite the fact that the imaginary components of the field are initialised randomly. However, when $\beta \lesssim 0.5$, they disagree. Moreover, the result from the cold start agrees with the one obtained in the world line formulation. One is therefore led to conclude that when $\mu = 0$ the imaginary components ϕ^{I} are driven to zero (more precisely, to a constant value) at large β but are not constrained at small β. In other words, the drift terms are not capable of restoring the reality of the dynamics. It is tempting to relate this to being in (or close to) the disordered phase. We note that it cannot be understood from the classical fixed point structure, since this is independent of β. We also remark that the dynamics at small β resembles Langevin dynamics with complex noise ($N_{\mathrm{I}} > 0$) [58], where the trajectories are kept in the complexified field space by the stochastic kicks on ϕ^{I} (rather than by the drift terms, as is the case here).

In terms of the distribution $P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}]$, these findings imply that $P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}] \sim e^{-S}\delta(\phi^{\mathrm{I}})$ at large β, but not at small β. This can be further investigated by

[2]The real components ϕ^{R} are taken from a Gaussian distribution always.

studying the width of the distribution in the imaginary direction,[3]

$$\left\langle \left(\Delta\phi^{\mathrm{I}}\right)^2 \right\rangle = \left\langle \frac{1}{\Omega}\sum_x \left(\phi_x^{\mathrm{I}}\right)^2 \right\rangle - \left\langle \frac{1}{\Omega}\sum_x \phi_x^{\mathrm{I}} \right\rangle^2 . \tag{3.1}$$

When $\mu = 0$ the width should vanish, while when turning on μ one may expect it to increase smoothly. The results are shown in Fig. 3.3. For the larger β values this is exactly what is observed: the width increases smoothly from zero. For the smaller β values, however, we observe that the width is nonzero even when $\mu = 0$ (when using a hot start), and remains large for nonzero μ. At larger values of μ the width is driven again towards zero and agreement with the world line results improves, see Fig. 3.2. We remark here that it is possible that different distributions (with different widths) yield the same result for observables. This is what is theoretically expected in the presence of complex noise ($N_{\mathrm{I}} \geq 0$) [58] and can be seen analytically in Gaussian models with complex noise, where a continuous family of distributions $P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}; N_{\mathrm{I}}]$ all yield the same result for observables, independent of N_{I}, even though the width of these distributions is nonzero and increases with N_{I} [75]. In the case we study here, however, we find that the failure of complex Langevin dynamics in the disordered phase is correlated with the spread of the distribution $P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}]$ in the noncompact direction. We conclude that a relatively narrow distribution, with a smoothly increasing width, is required. We note again that this resembles observations made in simulations of non-Gaussian models with complex noise [58, 76].

To investigate the interplay between (the width of) the distribution and observables, we express expectation values as

$$\langle A[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}] \rangle = \frac{1}{Z} \int D\phi^{\mathrm{R}} D\phi^{\mathrm{I}}\, P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}] A[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}], \tag{3.2}$$

[3]The mean value $\langle \phi^{\mathrm{I}} \rangle = 0$; in the large β phase, this requires averaging over a large number of initial conditions.

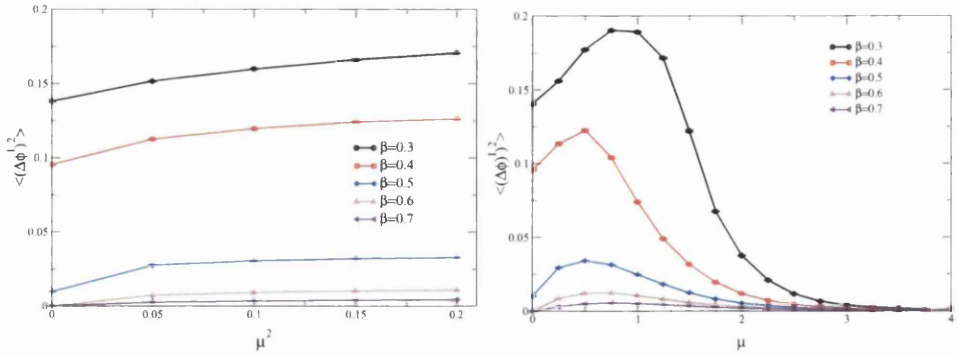


Figure 3.3: Width of the distribution $P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}]$ in the imaginary direction for various values of β as a function of μ^2 on a 10^3 lattice (left) and, for larger μ, as a function of μ on a 8^3 lattice (right).

with

$$Z = \int D\phi^{\mathrm{R}} D\phi^{\mathrm{I}}\, P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}]. \tag{3.3}$$

In general the operator A is not required to be holomorphic, since this will allow more insight in properties of the distribution.[4] The distribution of an operator A can then be defined according to

$$\langle A \rangle = \int dA\, P(A) A = \frac{1}{Z} \int D\phi^{\mathrm{R}} D\phi^{\mathrm{I}}\, P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}] A[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}], \tag{3.4}$$

where

$$P(A) = \frac{1}{Z} \int D\phi^{\mathrm{R}} D\phi^{\mathrm{I}}\, P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}] \delta(A - A[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}]), \tag{3.5}$$

with the normalization

$$\int dA\, P(A) = 1. \tag{3.6}$$

Distributions $P(A)$ can be constructed numerically, by sampling A from configurations generated by complex Langevin dynamics.

The distribution for the action density is shown in Fig. 3.4, comparing again a hot and cold start at $\mu = 0$. This figure supports the earlier claim

[4] Of course only holomorphic functions correspond to observables in the original theory.

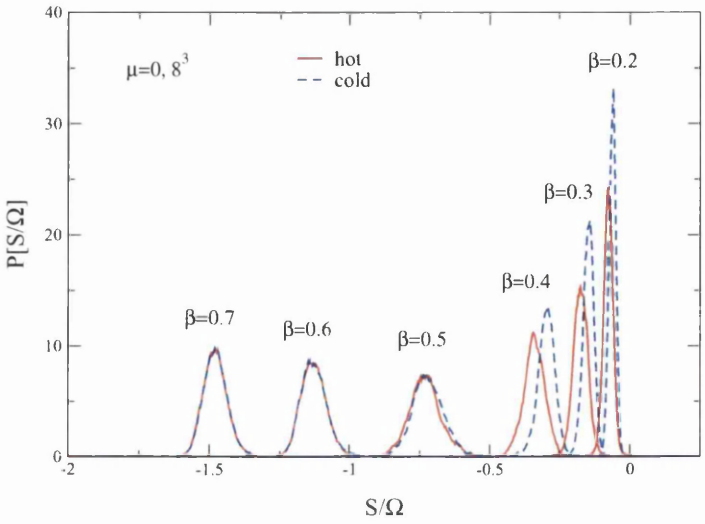
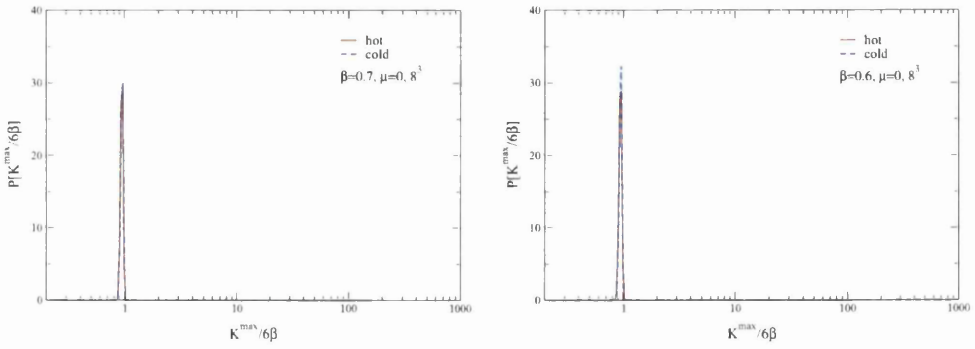


Figure 3.4: Distribution of action density S/Ω for various values of β at $\mu = 0$ on a 8^3 lattice, using a hot and a cold start.

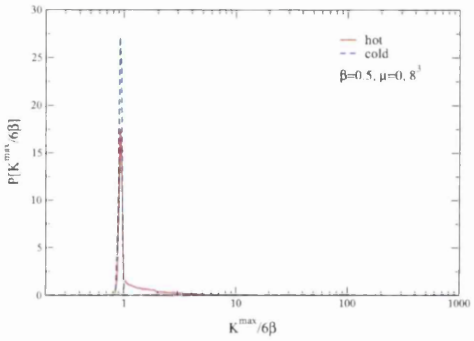
that real and complex Langevin match at larger β but fail at smaller β. However, the reason for failure is somewhat subtle. Naïvely, one might expect a large "tail" caused by excursions in the complexified field space to affect the expectation value but this does not appear to happen. Instead we find that the entire distribution is shifted and becomes only slightly wider at $\beta \lesssim 0.5$ when the hot start is used.

Finally, the observed difference at large and small β also appears prominently in the actual dynamics, i.e. in the drift terms. We have analyzed the maximal force $K^{\max}$ appearing in the adaptive stepsize algorithm. In the case of real Langevin dynamics, the drift terms are limited by an upper bound of $K^{\max} \leq 6\beta$. In the complexified space there is no upper limit and the drift terms can in principle become several orders of magnitude larger [57]. The distribution of $K^{\max}$ is plotted at $\mu = 0$ with hot and cold starts in Figure 3.5. In the large β phase, the distributions appear identical, with $K^{\max} \leq 6\beta$. This is consistent with the conclusion reached above. However, one should investigate whether the relevance of the extra drift term from a

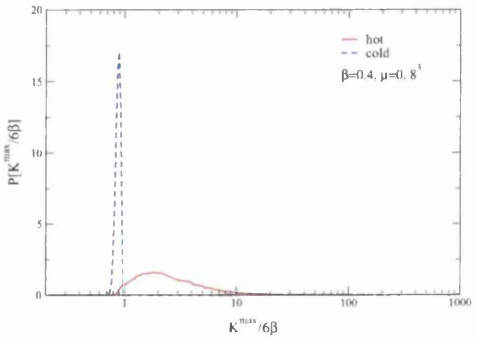


(a) $\beta = 0.7$

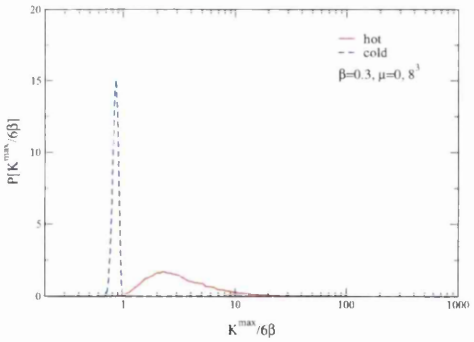
(b) $\beta = 0.6$



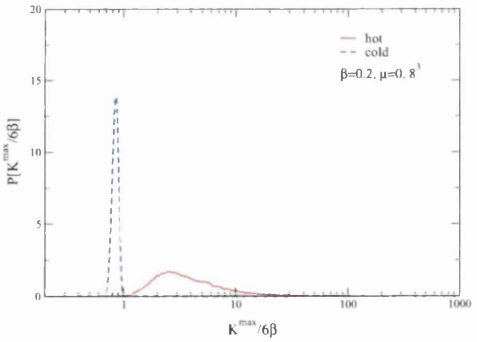
(c) $\beta = 0.5$



(d) $\beta = 0.4$

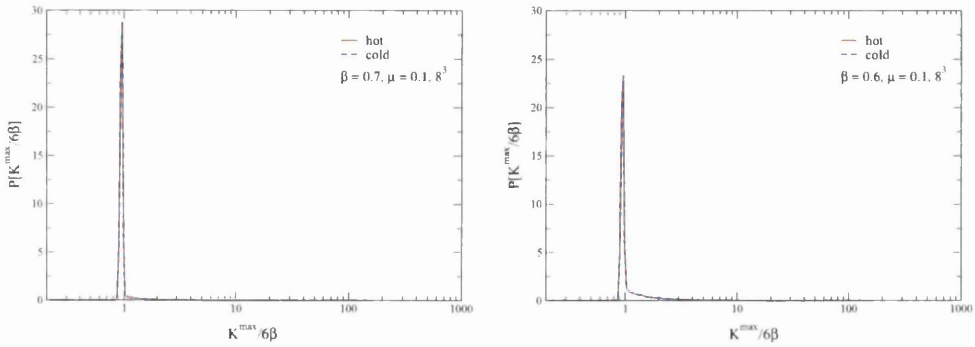


(e) $\beta = 0.3$



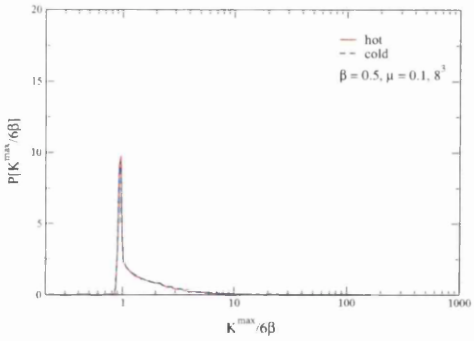
(f) $\beta = 0.2$

Figure 3.5: Distribution of $K^{\max}/(6\beta)$ at $\mu = 0$ on a 8^3 lattice using a hot and a cold start.

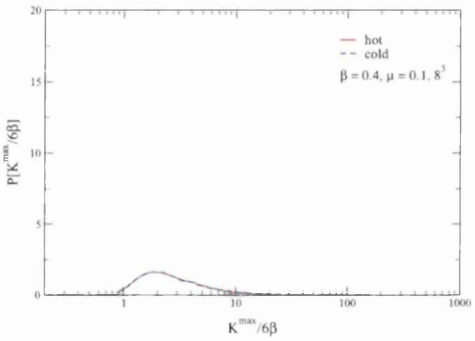


(a) $\beta = 0.7$

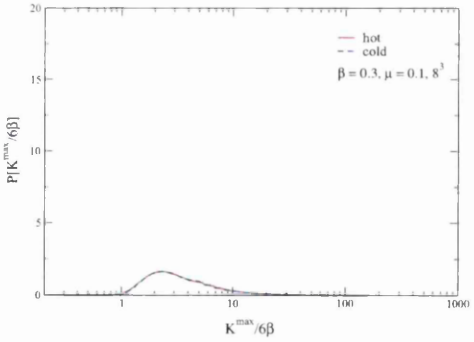
(b) $\beta = 0.6$



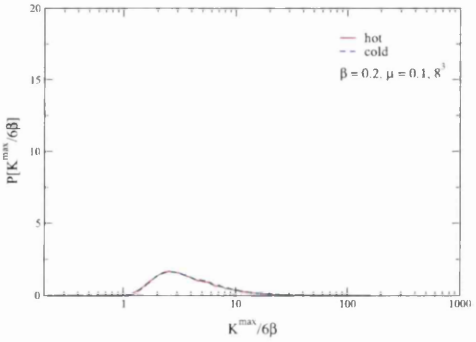
(c) $\beta = 0.5$



(d) $\beta = 0.4$



(e) $\beta = 0.3$



(f) $\beta = 0.2$

Figure 3.6: As in the previous figure, for $\mu = 0.1$

kernel would make a difference to this conclusion. In the low β phase the distributions are dramatically different: in the complexified dynamics, triggered by the hot start, much larger forces appear. The distributions are no longer peaked but very broad with a long tail (note the horizontal logarithmic scale). At $\beta = 0.5$ we observe interesting crossover behaviour: both the peaked distribution bounded by $K^{\max} = 6\beta$ and a decaying "tail" characteristic of small β distributions appear.

To study the two possible distributions of $K^{\max}$ further, we show in Fig. 3.6 the same results but now with $\mu = 0.1$. In this case the hot and cold starts yield identical distributions, since both simulations are complexified due to the nonzero chemical potential. The striking difference between the distributions at large and small β is still present. At large β the force can occasionally be large, making the use of an adaptive stepsize necessary. However, the typical value is still determined by the maximal value for real Langevin dynamics, i.e. 6β. At small β this part of the distribution is completely gone and is replaced by a broad distribution at much larger $K^{\max}$ values. Again at $\beta = 0.5$ we observe crossover behaviour with both features present. These results are qualitatively the same on larger volumes.

Let us summarize the findings of this section. Complex Langevin dynamics works well at large β in the ordered phase. The distribution $P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}]$ in the complexified field space is relatively narrow in the noncompact direction and Langevin simulations started with hot and cold initial conditions agree. The drift terms do occasionally become large but the typical size is set by the maximal value for real Langevin evolution. At small β, in or close to the disordered phase, the distribution is much wider in the ϕ^{I} direction. Typical drift terms are much larger, with a wide spread in the distribution. At $\mu = 0$ complexified dynamics does not reduce to real dynamics. There is a strong correlation with the phase the theory is in (see Fig. 3.2), but not with the sign problem, since these observations also hold at $\mu = 0$ and are independent of the lattice volume. Moreover, for the lattice volumes we consider the

sign problem is not severe. We emphasize that a firm conclusion can only be drawn after all the findings presented above are combined consistently, while the observation of e.g. large drift terms or a large width by itself would clearly be insufficient.

3.6 Conclusion

We have studied the applicability of complex Langevin dynamics to simulate field theories with a complex action due to a finite chemical potential, in the case of the three-dimensional XY model. Using analytical continuation from imaginary chemical potential and comparison with the world line formulation we found that complex Langevin dynamics yields reliable results at larger β but fails when $\beta \lesssim 0.5$ at small chemical potential. We established that the region of failure is strongly correlated with the part of the phase diagram which corresponds to the disordered phase. We have verified that these conclusions do not depend on the lattice volume. Failure at small β values was also observed a long time ago in the case of SU(3) field theory in the presence of static charges [47].

Due to the use of an adaptive stepsize algorithm no runaways or instabilities have been observed. The results we found in the disordered phase are therefore interpreted as convergence to the wrong result. To analyze this, we have studied properties of the dynamics and field distributions in the complexified field space. For the smaller β values, we found that complexified dynamics does not reduce to real dynamics when $\mu = 0$. Furthermore, for the system sizes and parameter values we used, the sign problem is not severe. We conclude therefore that the failure is not due to the presence of the sign problem, but rather due to an incorrect exploration of the complexified field space by the Langevin evolution. The forces appearing in the stochastic process behave very differently at large and small β. Interestingly, in the crossover region at $\beta \approx 0.5$, the dynamics shows a combination of large and

small β characteristics.

Chapter 4

Component representation

4.1 Changing representation

Many theories have different but equivalent representations and results should in principle be independent of this choice. As was discussed in Chapter 3 with the XY model, some theories can be transformed and written in terms of new variables in such a way that the sign problem is removed. However, this relies on the existence of such dual variables and in general this is not the case. What is often possible, however, is to transform the degrees of freedom, which are typically elements of a Lie group, to those in a different representation.

The degrees of freedom in which the XY model was written in Chapter 3 were angular spin variables. However, there is the freedom to transform a polar coordinate into a rectangular coordinate by writing the spin φ in terms of components, $\phi_1 = \cos\varphi$, $\phi_2 = \sin\varphi$, so that

$$\phi_1 + i\phi_2 = e^{i\varphi}. \tag{4.1}$$

This comes with the constraint on the two variables,

$$\phi_1^2 + \phi_2^2 = 1. \tag{4.2}$$

All expressions can then be expanded and written in terms of the new component variables $\phi_{1,2}$ instead of the spin angle.

In this formalism, the action of the XY model is

$$S = -\beta \sum_{x,\nu} \cos(\varphi_x - \varphi_{x+\hat{\nu}} - i\mu\delta_{\nu,0}), \tag{4.3}$$

$$= -\frac{\beta}{2} \sum_{x,\nu} \phi_{a,x}\phi_{a,x+\hat{\nu}} \cosh(\mu\delta_{\nu,0}) - i \sinh(\mu\delta_{\nu,0})\epsilon_{a,b}\phi_{a,x}\phi_{b,x+\hat{\nu}}, \tag{4.4}$$

with $a = 1, 2$. Complexification is then done separately on each ϕ_a so that the drift terms in the Langevin equation are in the direction of each of the components,

$$K_{a,x} = -\left.\frac{\delta S}{\delta \phi_{a,x}}\right|_{\phi_{a,x} \to \phi^{\mathrm{R}}_{a,x} + i\phi^{\mathrm{I}}_{a,x}}. \tag{4.5}$$

Explicitly they read

$$K^{\mathrm{R}}_{a,x} = -\frac{\beta}{2} \sum_{\nu} \left[\cosh(\mu\delta_{\nu,0})(\phi^{\mathrm{R}}_{a,x+\hat{\nu}} + \phi^{\mathrm{R}}_{a,x-\hat{\nu}}) + \sinh(\mu\delta_{\nu,0})\epsilon_{ab}(\phi^{\mathrm{I}}_{b,x+\hat{\nu}} - \phi^{\mathrm{I}}_{b,x-\hat{\nu}})\right] \tag{4.6}$$

$$K^{\mathrm{I}}_{a,x} = -\frac{\beta}{2} \sum_{\nu} \left[\cosh(\mu\delta_{\nu,0})(\phi^{\mathrm{I}}_{a,x+\hat{\nu}} + \phi^{\mathrm{I}}_{a,x-\hat{\nu}}) - \sinh(\mu\delta_{\mu,0})\epsilon_{ab}(\phi^{\mathrm{R}}_{b,x+\hat{\nu}} - \phi^{\mathrm{R}}_{b,x-\hat{\nu}})\right]. \tag{4.7}$$

The Langevin equation therefore consists of updates for both the complexified variables

$$\phi^{\mathrm{R}}_{a,x}(n+1) = \phi^{\mathrm{R}}_{a,x} + \epsilon K^{\mathrm{R}}_{a,x} + \sqrt{\epsilon}\eta_{a,x} \tag{4.8}$$

$$\phi^{\mathrm{I}}_{a,x}(n+1) = \phi^{\mathrm{I}}_{a,x} + \epsilon K^{\mathrm{I}}_{a,x}. \tag{4.9}$$

The Langevin update does not preserve the constraint given by Eq. (4.2) and so it must be imposed to form the complete update.

4.2 Applying the constraint

After each update, the constraint $|\phi|^2 = \phi_1^2 + \phi_2^2 = 1$ will no longer hold. Therefore, the fields must be modified after each update so that the variables are put back onto the complexified circle. A rescaling procedure preserves the symmetry between both components,

$$\phi_1 + i\phi_2 \to \frac{\phi_1 + i\phi_2}{\sqrt{\phi_1^2 + \phi_2^2}}. \tag{4.10}$$

However, after complexification, the denominator will be complex,

$$\phi_1^2 + \phi_2^2 \to \phi_a^{\mathrm{R}\,2} - \phi_a^{\mathrm{I}\,2} + 2i\phi_a^{\mathrm{R}}\phi_a^{\mathrm{I}}. \tag{4.11}$$

In order to take a square root of a complex number, one must choose a branch cut. The usual choice for this is the negative real axis, which shall be used here. The principle square root of a complex number $z = re^{i\varphi}$ is then given by

$$\sqrt{z} = \sqrt{r}e^{i\varphi/2}, \tag{4.12}$$

with $-\pi \leq \varphi < \pi$. In component representation, $z = x + iy$, Eq. (4.12) is equivalent to

$$\sqrt{z} = \sqrt{\frac{r+x}{2}} + i\mathrm{sign}(y)\sqrt{\frac{r-x}{2}}. \tag{4.13}$$

Using this definition, the rescaling procedure to put the spin components back onto the complexified circle can be specified. Starting with the defini-

tions

$$a = \mathrm{Re}\,(\phi_1^2 + \phi_2^2) = \phi_a^{\mathrm{R}\,2} - \phi_a^{\mathrm{I}\,2}, \tag{4.14}$$

$$b = \mathrm{Im}\,(\phi_1^2 + \phi_2^2) = 2i\phi_a^{\mathrm{R}}\phi_a^{\mathrm{I}}, \tag{4.15}$$

$$r = \sqrt{a^2 + b^2}, \tag{4.16}$$

the scaling factor is

$$A = \sqrt{\phi_1^2 + \phi_2^2} = \sqrt{a + ib}$$

$$A^{\mathrm{R}} + iA^{\mathrm{I}} = \sqrt{\frac{r+a}{2}} \pm i\sqrt{\frac{r-a}{2}}. \tag{4.17}$$

After complexification, the rescaling of ϕ_a is then

$$\phi_a^{\mathrm{R}} \to \frac{A_{\mathrm{R}}\phi_a^{\mathrm{R}} + A_{\mathrm{I}}\phi^{\mathrm{I}}}{r}, \tag{4.18}$$

$$\phi_a^{\mathrm{I}} \to \frac{A_{\mathrm{I}}\phi_a^{\mathrm{R}} - A_{\mathrm{R}}\phi_a^{\mathrm{I}}}{r}. \tag{4.19}$$

This procedure therefore defines a general method for updating a component form of the spin variable and rescaling it back onto the complexified circle. The rescaling prescription suffers from a problem, however. If the components are one of the forms

$$\phi_1^{\mathrm{R}} = \phi_2^{\mathrm{I}}, \qquad \phi_1^{\mathrm{I}} = -\phi_2^{\mathrm{R}}, \tag{4.20}$$

$$\phi_1^{\mathrm{R}} = -\phi_2^{\mathrm{I}}, \qquad \phi_1^{\mathrm{I}} = \phi_2^{\mathrm{R}}, \tag{4.21}$$

then $A = 0$ meaning that the denominator vanishes and the rescaling procedure breaks down. By transforming back to the angular representation,

$$\phi_1^{\mathrm{R}} = \cos\varphi^{\mathrm{R}}\cosh\varphi^{\mathrm{I}}, \qquad \phi_1^{\mathrm{I}} = -\sin\varphi^{\mathrm{R}}\sinh\varphi^{\mathrm{I}}, \tag{4.22}$$

$$\phi_2^{\mathrm{R}} = \sin\varphi^{\mathrm{R}}\cosh\varphi^{\mathrm{I}}, \qquad \phi_2^{\mathrm{I}} = \cos\varphi^{\mathrm{R}}\cosh\varphi^{\mathrm{I}}, \tag{4.23}$$

the configurations above where $A = 0$ correspond to where the imaginary part of the angle φ^{I} diverges to infinity

$$\cosh \varphi^{\mathrm{I}} = \sinh \varphi^{\mathrm{I}}. \tag{4.24}$$

This is a problem in the angular representation too, although numerical simulations will break down due to the limitations of computers for storing large floating point numbers. The adaptive stepsize algorithm used in the XY model prevents these configurations from being generated by discouraging updates in this direction.

Although such configurations are permissible in this formulation, they do not cause such numerical problems since the scaling factor A never exactly vanishes, it can be small, however. This has the effect of sending the components further to infinity in order for the constraint to be satisfied. This means that the component representation can explore more of the configuration space than is possible with the spin representation.

4.3 Results

Comparing results from the spin formulation with the new component representation, it can be seen in Figure 4.1 that although agreement with the correct result (using the world line formalism) is better than with the spin formulation, unfortunately it still does not give correct results. This can be seen clearly when plotted as a colour map, comparing the expectation value for $\mathrm{Re}\,\langle S \rangle$ using complex Langevin and world line methods. The Silver Blaze feature is again not present with the Langevin data, resulting in the strongest discrepancy in this region.

Examining the histograms of certain observables can provide some insight into the space that is being explored by the simulation.

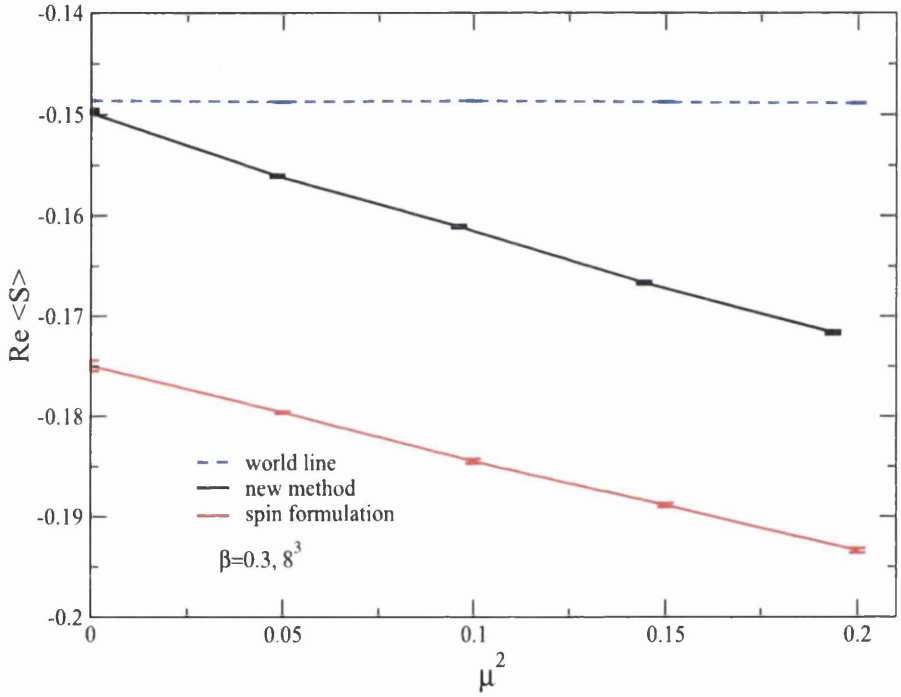


Figure 4.1: Plot of $\mathrm{Re}\,\langle S\rangle/\Omega$ on 8^3 lattice at $\beta = 0.3$.

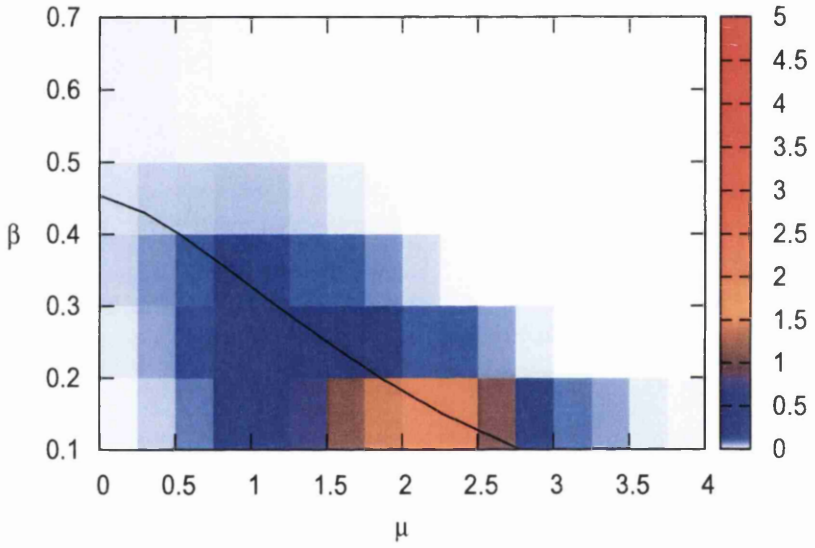


Figure 4.2: Colour map of relative error using component representation. Compare with the equivalent plot using the spin formulation.

4.4 Vortex loops

Vortices are known to play an important part in the Kosterlitz-Thouless phase transition of the 2d XY model. In the 3d XY model, vortices are not confined to the plane, but instead form non-trivial loops through the lattice volume. There is a possibility that different representations of the XY model, in spin or component form, will give rise to different vortex behaviors.

A vortex is loosely defined to be the winding number of the angles around a fixed point in a field ϕ, which can be written as

$$v_x = \frac{1}{2\pi} \oint_C d\phi \tag{4.25}$$

around a closed contour C. Since the field ϕ is required to be continuous, the integral will be an integer multiple of 2π. The winding number is therefore v_x and is an integer. The points around which v_x is positive are called vortices, whilst those with negative v_x are called anti-vortices. A more precise definition, in the case of a spin field on a lattice, is

$$\begin{aligned} v_{x,\mu\nu} = \frac{1}{2\pi} \Big(& [\phi_x - \phi_{x+\hat{\mu}}; 2\pi] + [\phi_{x+\hat{\mu}} - \phi_{x+\hat{\mu}+\hat{\nu}}; 2\pi] + \\ & [\phi_{x+\hat{\nu}} - \phi_{x+\hat{\mu}+\hat{\nu}}; 2\pi] + [\phi_{x+\hat{\nu}} - \phi_x; 2\pi] \Big) , \end{aligned} \tag{4.26}$$

where taking $[x; y]$ maps the real number x to the interval $[-y/2, y/2)$. Therefore, the winding number v_x, which is around the point $x + \hat{\mu}/2 + \hat{\nu}/2$ can be one of $-1, 0, 1$.

By considering a sum of neighbouring winding numbers, it can be seen that the shared link will cancel,

$$v_{x,\mu\nu} + v_{x+\hat{\mu},\mu\nu} = \Big([\phi_{x+\hat{\mu}} - \phi_{x+\hat{\mu}+\hat{\nu}}; 2\pi] + [\phi_{x+\hat{\mu}+\hat{\nu}} - \phi_{x+\hat{\mu}}; 2\pi] \Big) + \dots . \tag{4.27}$$

Therefore it follows that the sum over the lattice volume of v_x will vanish and hence there are always an equal number of vortices and anti-vortices.

In the case of the 3d XY model, vortices form closed loops which wind through the volume. The procedure to trace all the loops in a given configuration $\{\phi_x\}$ begins by observing that the lattice volume can then be considered to consist of face-centred cubes, with each face of the cube (a plaquette) having a winding number. Note the relationship between the "outward" (v^+) and "inward" (v^-) facing loops $v^-_{x,\mu\nu} = -v^+_{x,\mu\nu}$. Each of the plaquettes is initially marked as "unread" to avoid the possibility of multiple counting. Starting from a particular cube, labelled at site x, compute the available exit plaquettes, which are those plaquettes from $v^-_{x,01}, v^-_{x,02}, v^-_{x,12}, v^+_{x+\hat{2},01}, v^+_{x+\hat{1},02}, v^+_{x+\hat{0},12}$ for which $v = +1$ and are unread. If there is only a single available exit, then tag this plaquette as "read" and repeat the process on the cube in that direction until the loop arrives back at the original site. If there is more than one available exit then there is an ambiguity and the tracing procedure must make a choice of how to deal with these situations. The two most commonly used procedures are either to choose the exit which maximises loop length or to make the choice at random, which is the scheme used here. If there are no exits from the cube then an error has occurred since all vortex loops must eventually close. By moving systematically through the lattice, all the plaquettes will eventually be marked as "read" which signifies that all loops have been traced.

Note that this procedure is identical in the component formulation, since the angle can be computed from the components,

$$\varphi_x = \begin{cases} \arctan(\phi^{\mathrm{I}}/\phi^{\mathrm{R}}) & \text{if } \phi^{\mathrm{R}} > 0 \\ \arctan(\phi^{\mathrm{I}}/\phi^{\mathrm{R}}) + \pi & \text{if } \phi^{\mathrm{I}} \geq 0, \phi^{\mathrm{R}} < 0 \\ \arctan(\phi^{\mathrm{I}}/\phi^{\mathrm{R}}) - \pi & \text{if } \phi^{\mathrm{I}} \leq 0, \phi^{\mathrm{R}} < 0 \\ +\pi/2 & \text{if } \phi^{\mathrm{I}} > 0, \phi^{\mathrm{R}} = 0 \\ -\pi/2 & \text{if } \phi^{\mathrm{I}} < 0, \phi^{\mathrm{R}} = 0 \\ \text{undefined} & \text{if } \phi^{\mathrm{I}} = 0, \phi^{\mathrm{R}} = 0 \end{cases} \tag{4.28}$$

length	number
4	12
6	4
8	2
10	3
12	2
14	2
18	2
32	1
58	1
74	1
148	1
1458	1

Table 4.1: Number of plaquettes in a typical configuration at $\beta = 0.3$, $\mu = 0.5$ on a 12^3 lattice.

which is provided by the `atan2` function in most computer programming systems. Note that in both the spin and component representations, the imaginary part of the angle ϕ^{I} is ignored in the process of computing vortex loops. Note that in the world line formalism the angles cannot be recovered, since they are integrated out and replaced with new degrees of freedom.

The ordered and disordered phases are expected to display different vortex behaviours. In an ordered phase, where neighbouring angles are closely aligned, there will be few plaquettes with non-zero winding numbers and hence there will be few vortex loops. The disordered phase is expected to have more vortex loops since the neighbouring spins are not closely correlated and therefore non-zero winding numbers are more likely.

An example configuration, generated at $\beta = 0.3, \mu = 0.5$ on a 12^3 lattice, has a single large vortex loop and several short loops, shown in Table 4.1. To get any meaningful information about vortex loops, it is necessary to average over many configurations, as with any other observable. Two quantities may be computed, the average winding number $\langle v^{+}_{x,\mu\nu} \rangle / (3\Omega)$ and the average loop

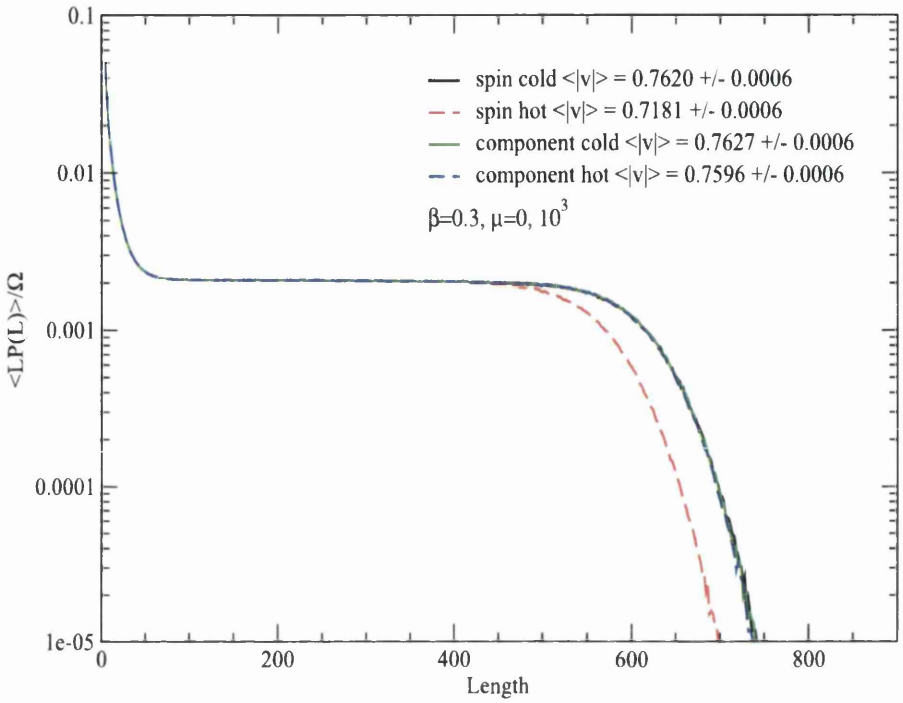


Figure 4.3: Histogram of fraction of volume taken by vortices as a function of length at $\mu = 0$, with hot and cold starts and both spin and component representations.

length. However, since it is important to distinguish between a small number of long loops and a large number of short loops (which may have a similar mean length), a histogram of vortex loop length gives more information.

The distribution of the fraction of lattice volume consumed by vortices as a function of length is shown in Figure 4.3, computed using hot and cold starts with both spin and component representation. This is a useful quantity because it shows clearly that longer, less frequent vortex loops are just as important as the more frequent but shorter loops. Up to a certain point the contribution is $LP(L)/\Omega$ is flat and consistent for both hot/cold starts with both representations. However, the erroneous data for hot start spin representation decays earlier than the cold start and component representa-

tion data. This means that the missing vortices in the spin representation complex Langevin simulations come exclusively from the long vortices that wind around the volume.

The discrepancy in vortices has a direct impact on other expectation values, importantly on the value of $\cos(\varphi_x - \varphi_{x+\hat{\nu}})$. In order for a plaquette to have a non-zero winding number, on at least one of the links the difference must be greater than π,

$$|\varphi_x - \varphi_{x+\hat{\nu}}| \geq \pi. \tag{4.29}$$

This will be evident in a histogram of the sampled values of $\cos(\varphi_x - \varphi_{x+\hat{\nu}})$. Plaquettes which have a non-zero winding number will have links which contribute towards $\cos\Delta\varphi \sim -1$, whereas if the winding number is zero the links will be contributing towards $\cos\Delta\varphi \sim +1$. This discrepancy can be seen in Figure 4.4. In the ordered phase, at high β, the histograms of Re $\cos(\Delta\varphi_x)$ match, which is not unexpected considering that other observables also agree. However, it is in the disordered phase at low β where the discrepancy is evident. The spin representation samples a much wider range of values for Re $\cos(\Delta\varphi_x)$, as can be seen from the slowly decaying tails. As as result, the peak at -1 is lower which indicates that there are fewer vortices present.

4.5 Conclusion

The failure of the complex Langevin method with the XY model occurs even when $\mu = 0$ and the sign problem is absent. By using hot initial conditions, in which the variables are spread in the complex plane, the real and complex dynamics can be compared. In the ordered phase, where the complex Langevin results are correct, the variables are driven back to the real axis after thermalisation. However, in the disordered phase the dynamics fails to attract the variables back to the real axis. This is in contrast to the Bose gas, for which complex Langevin dynamics was found to give correct results in all parts of the phase diagram. This difference is surprising when the

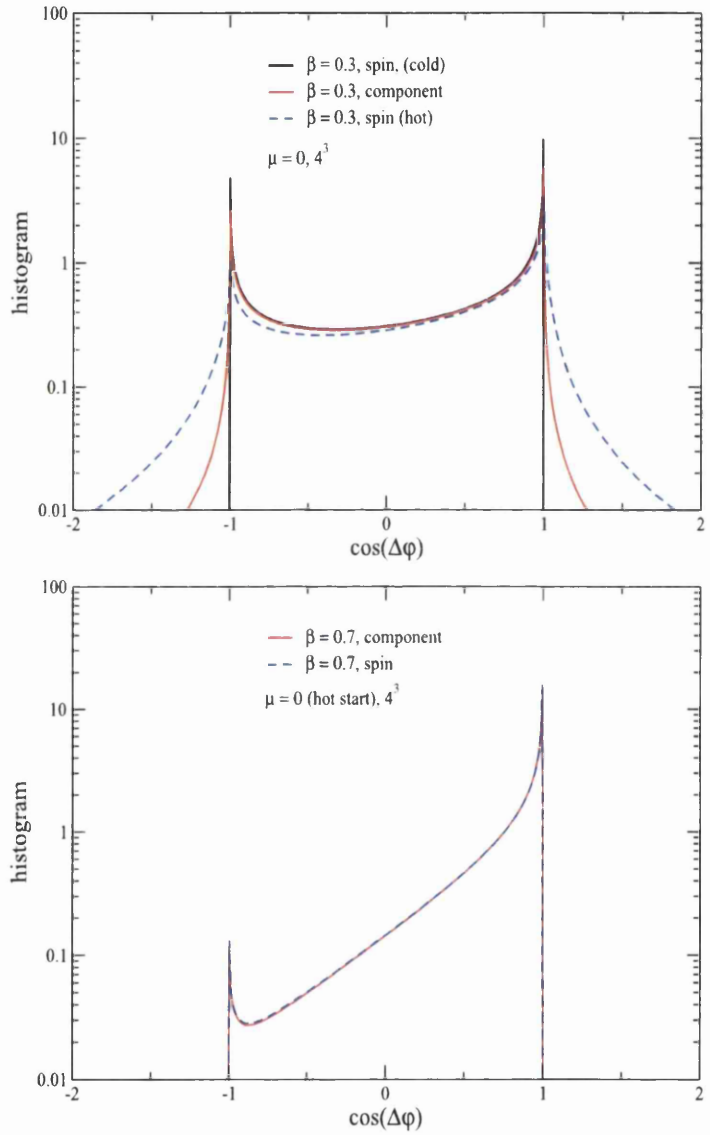


Figure 4.4: Histograms of Re $\cos(\Delta\phi_x)$, comparing spin and component representation, at high and low β on a 4^3 lattice.

actions of the two theories are examined by rewriting the spin variables of the XY model in terms of components. Doing this it can be seen that the XY model is merely the limit of the Bose gas in which $\lambda \to \infty$ and the mass is chosen to be $m^2 = -2d = -6$. Unfortunately, it is not possible to realise the XY model in this way, because of the difficulty in taking the limit in λ. By rewriting the XY model in terms of components, which mirror the real and imaginary components in the Bose gas, and adding the constraint that $\phi_1^2 + \phi_2^2 = 1$, the similarity in the actions is apparent.

The results that the component representation gives are closer to the correct results, found from the world line method, than from using the spin representation, but are still not exactly correct in the low-β phase. The difference between the two representations in the disordered phase can be seen by examining the vortices, in particular the distribution of vortices as a function of length. The evidence from histograms of both vortex loop length and the more simple measure of $\cos(\Delta_\nu \varphi)$ show that there are slightly too few vortices in the spin representation. The vortices that are in that sense missing are those that are very long and wrap around the whole volume.

Chapter 5

Justification and criteria for correctness

The contents of this chapter first appeared in [58], and then in Ref. [77] of which I am an author.

5.1 Introduction

The problem of convergence to the wrong limit, where the simulation gives well defined but incorrect results, is not limited to the case of the XY model in Chapter 3, but has been found since the early complex Langevin studies of the 1980s. Incorrect convergence remains the main issue preventing complex Langevin dynamics being widely used for complex action problems. In this chapter we cover a formal justification for complex Langevin dynamics and identify some points at which it might fail. By studying the long time evolution of observables with respect to real and complex measures, we derive a criterion which must be satisfied in order for correct results to be obtained [58, 77].

The central object of interest is the expectation value of a particular

observable, given by

$$\langle O \rangle = \frac{\int O(x) e^{-S(x)} dx}{\int e^{-S(x)} dx}, \tag{5.1}$$

where for notational simplicity we use a single real degree of freedom, x. The action $S(x)$ is complex, preventing a probability interpretation of the measure and ruling out methods based on importance sampling. The complex Langevin equation is

$$\frac{dx}{dt} = K_x + \sqrt{N_{\mathrm{R}}}\eta_{\mathrm{R}}, \qquad \frac{dy}{dt} = K_y + \sqrt{N_{\mathrm{I}}}\eta_{\mathrm{I}}, \tag{5.2}$$

where the real variable is analytically continued as $x \to z = x + iy$ and t is Langevin time. The drift terms are given by

$$K_x = -\mathrm{Re}\left.\frac{dS(x)}{dx}\right|_{x \to x+iy}, \qquad K_y = -\mathrm{Im}\left.\frac{dS(x)}{dx}\right|_{x \to x+iy}, \tag{5.3}$$

and the two noise terms η_{R}, η_{I} are independent Gaussian random numbers with variance 2 and normalisation $N_{\mathrm{I}} \geq 0$ and $N_{\mathrm{R}} - N_{\mathrm{I}} = 1$. A numerical simulation can then be implemented by integrating these equations to large times $t \to \infty$.

By Itô calculus, if f is a twice differentiable function and $z(t)$ is a solution of Eq. (5.2) then

$$\frac{d}{dt}\langle f(z) \rangle = \langle Lf(z) \rangle, \tag{5.4}$$

where L is the operator

$$L = \left[N_{\mathrm{R}}\nabla_x + K_x\right]\nabla_x + \left[N_{\mathrm{I}}\nabla_y + K_y\right]\nabla_y, \tag{5.5}$$

and $\langle f \rangle$ denotes a noise average with respect to the stochastic process.

5.2 Formal justification

The dynamics resulting from solving the complex Langevin equation is described by a dual Fokker-Planck equation for the evolution of the probability density $P(x, y; t)$,

$$\frac{\partial}{\partial t} P(x, y; t) = L^T P(x, y; t), \tag{5.6}$$

with the operator

$$L^T = \nabla_x [N_{\mathrm{R}} \nabla_x - K_x] + \nabla_y [N_{\mathrm{I}} \nabla_y - K_y], \tag{5.7}$$

the formal transpose of Eq. (5.5) with respect to the bilinear pairing

$$\langle P, f \rangle = \int f(x, y) P(x, y) dx dy, \tag{5.8}$$

so that

$$\langle P, Lf \rangle = \langle L^T P, f \rangle. \tag{5.9}$$

To understand the time evolution of the real density $P(x, y; t)$ one must also examine the evolution of the complex density $\rho(x; t)$, determined by

$$\frac{\partial}{\partial t} \rho(x; t) = L_0^T \rho(x; t). \tag{5.10}$$

Here, the complex Fokker-Planck operator L_0^T is

$$L_0^T = \nabla_x [\nabla_x + \nabla_x S(x)]. \tag{5.11}$$

This equation has $\rho(x; \infty) \propto \exp[-S(x)]$ as a stationary solution, which is expected to be unique. Numerical studies (where feasible) of Eq. (5.10) confirm this to be true; in fact, convergence to this distribution seems exponentially fast.

Expectation values with respect to the two densities can now be defined

as

$$\langle O \rangle_{P(t)} = \frac{\int O(x+iy) P(x,y;t) dx dy}{\int P(x,y;t) dx dy}, \tag{5.12}$$

$$\langle O \rangle_{\rho(t)} = \frac{\int O(x) \rho(x;t) dx}{\int \rho(x;t) dx}. \tag{5.13}$$

The result that one would like to show is

$$\langle O \rangle_{P(t)} = \langle O \rangle_{\rho(t)}, \tag{5.14}$$

if the initial conditions $\langle O \rangle_{P(0)} = \langle O \rangle_{\rho(0)}$ match, which is assured provided

$$P(x,y;0) = \rho(x;0)\delta(y). \tag{5.15}$$

One expects the dependence on the initial conditions to vanish in the limit $t \to \infty$ by ergodicity.

To establish a connection between the expectation values with respect to P and ρ, one moves the time evolution from the densities to the observables. Since we are only interested in functions of $z = x + iy$ (holomorphic functions), we may act with the Langevin operator

$$\tilde{L} = [\nabla_z - (\nabla_z S(z))]\nabla_z.$$

The Cauchy-Riemann equations imply that $\nabla_y = i\nabla_x$ and therefore the action of L and $\tilde{L}$ on holomorphic functions agree. Since we only consider holomorphic observables here, L and $\tilde{L}$ may be used interchangeably.

We now use $\tilde{L}$ to evolve observables according to the equation

$$\frac{\partial}{\partial t} O(z;t) = \tilde{L} O(z;t), \tag{5.16}$$

which is formally solved by

$$O(z;t) = \exp[t\tilde{L}]O(z).$$

This requires that the operators $L, \tilde{L}$ and their transposes may be exponentiated. Although a general mathematical answer to this question remains elusive, numerical results show this to be true in all cases considered.

To examine the evolution we define the function

$$F(t,\tau) = \int P(x,y;t-\tau)O(x+iy;\tau)dxdy, \tag{5.17}$$

and observe that $F(t,\tau)$ interpolates between the two expectation values:

$$F(t,0) = \langle O \rangle_{P(t)}, \tag{5.18}$$

$$F(t,t) = \langle O \rangle_{\rho(t)}. \tag{5.19}$$

The first can be seen easily, while the second makes use of the initial conditions and

$$\begin{aligned} F(t,t) &= \int P(x,y;0)\left(e^{tL}O\right)(x+iy;0)dxdy \\ &= \int \rho(x;0)\left(e^{tL_0}O\right)(x;0)dx = \int O(x;0)\left(e^{tL_0^T}\rho\right)(x;0)dx \\ &= \langle O \rangle_{\rho(t)}, \end{aligned} \tag{5.20}$$

where it is only necessary to assume that integration by parts in x does not produce any boundary terms.

The desired result of Eq. (5.14) follows if $F(t,\tau)$ is independent of τ. To

check this, we need the τ derivative to vanish,

$$\frac{\partial}{\partial\tau}F(t,\tau) = -\int \left(L^T P(x,y;t-\tau)\right) O(x+iy;\tau)dxdy + \int P(x,y;t-\tau)LO(x+iy;\tau)dxdy. \tag{5.21}$$

Integration by parts, *if applicable without boundary terms at infinity*, then shows that the two terms cancel and therefore $F(t,\tau)$ is independent of τ, irrespective of $N_{\rm I}$.

This is therefore a point at which the formal argument might fail: if the decay of the product

$$P(x,y;t-\tau)O(x+iy;\tau)$$

and its derivatives is insufficient for integration by parts without boundary terms.

In Ref. [77] a study of the U(1) one-link model found that the τ-derivative is largest at $\tau = 0$. This motivates the superficially weaker condition

$$\lim_{t\to\infty}\frac{\partial}{\partial t}F(t,\tau)\Big|_{\tau=0} = 0. \tag{5.22}$$

This modification to the condition is still a sufficient criterion for correctness, provided it holds for a sufficiently large set of suitably chosen observables. Taking the limit $t \to \infty$ in Eq. (5.21) causes the first contribution to vanish because of the equilibrium condition, $L^T P(x,y;\infty) = 0$. Therefore, the criterion for correctness reduces to

$$E_O \equiv \int P(x,y;\infty)LO(x+iy,0)dxdy = \langle LO\rangle = 0. \tag{5.23}$$

This is fairly simple to check for a given observable, but it is in fact a strong statement since it must hold *for all observables*. Therefore, Eq. (5.23) really represents an infinite tower of identities which must all be satisfied. Since the criterion must be satisfied for all observables, if there is at least one

observable for which it is violated then the justification definitely breaks down. Hence, in practice the criteria can be checked for a small number of observables which still yields a necessary criterion [77].

5.3 Conclusion

Complex Langevin dynamics is a general procedure that can be applied where the sign problem prevents the use of importance sampling. The vexing problem of convergence to the wrong limit, which at present cannot be cured, remains the major barrier preventing the method from general usage.

Although the formal proofs that Langevin converges correctly break down when the action is complex, the arguments described in this chapter set out a justification of the method. The conclusion is that, provided the various integration by parts can be done without boundary terms, expectation values from the complex Langevin process are correct. This is an important result, because it shows that provided certain criteria are met, complex Langevin dynamics can be relied upon to give correct results. The justification goes further, because it shows where to look in simulation data to decide whether the criteria have been met. By computing the consistency condition $\langle LO \rangle \stackrel{?}{=} 0$ and the decay rate of histograms it is possible to infer whether the boundary terms are important. Therefore these are self consistent checks and are independent of other methods.

Chapter 6

SU(3) spin model

6.1 Introduction

In this chapter we consider the three-dimensional SU(3) spin model at nonzero density, an effective Polyakov loop model which follows from the QCD Lagrangian in a combined strong-coupling and heavy-quark expansion and one of the first QCD-related models addressed with complex Langevin dynamics [78, 79]. Our reason to revisit this model is partly due to the recent discussion of Gattringer, who showed how a reformulation in terms of fluxes eliminates the sign problem [37]. Moreover, given that our understanding of complex Langevin dynamics has steadily improved in the past years [?,51–56,80–83], we consider it worthwhile to reconsider the model and apply recently developed tools [80,81,83] to assess the applicability of complex Langevin dynamics in detail, something that was not undertaken in the classic papers [78,79].

This chapter is organized as follows. In the next section we introduce the SU(3) model and summarize some basic results at finite density. The complex Langevin equations are given in Section 6.3. Besides the standard lowest-order discretization, we also describe a higher-order algorithm to eliminate leading stepsize corrections [84]. In Section 6.4 we discuss our current under-

standing of the applicability of complex Langevin dynamics at finite density and review the various ways in which the outcome of a complex Langevin process can be assessed, in particular when the exact result is not available. Section 6.5 constitutes the main part of the chapter. Here we present a variety of numerical results assessing the applicability of complex Langevin dynamics in this model, both in the disordered and the ordered phase. We also demonstrate that the higher-order algorithm eliminates most of the stepsize dependence. Conclusions are drawn in Section 6.7. The higher-order algorithm is discussed in some more detail in Appendix A, while Appendix B can be used to scrutinize the stepsize dependence and criteria for correctness.

6.2 SU(3) spin model

We consider the three-dimensional SU(3) spin model at nonzero chemical potential, with the action [78]

$$S = S_B + S_F, \tag{6.1}$$

where

$$S_B = -\beta \sum_x \sum_{\nu=1}^{3} \left(\mathrm{Tr}\, U_x \mathrm{Tr}\, U^\dagger_{x+\hat{\nu}} + \mathrm{Tr}\, U^\dagger_x \mathrm{Tr}\, U_{x+\hat{\nu}} \right), \tag{6.2}$$

$$S_F = -h \sum_x \left(e^{\mu} \mathrm{Tr}\, U_x + e^{-\mu} \mathrm{Tr}\, U^\dagger_x \right). \tag{6.3}$$

The model can be thought of as an effective dimensionally reduced version of QCD, where $\mathrm{Tr}\, U_x$ represents the trace of the Polyakov loop; the U_x's are SU(3) matrices living on a three-dimensional lattice (we use periodic boundary conditions). The first term then represents the gluon contribution with effective coupling β, while the second term represents heavy quarks, with coupling h. Chemical potential favours quarks over anti-quarks, resulting in a complex action, $S_F^*(\mu) = S_F(-\mu^*)$. The fermion term is a simplified version

of the contribution derived in the heavy dense limit [54]. The partition function,

$$Z = e^{-\Omega f} = \int \prod_x dU_x \, e^{-S}, \tag{6.4}$$

is even in μ due to charge conjugation invariance. Here f denotes the free energy density and Ω is the three-dimensional volume.

The phase structure of this theory has been studied in Refs. [78,79], using both complex Langevin dynamics and mean-field theory. Recently it has also been investigated using a reformulation of the theory which is sign-problem free [38]. For small h, the theory has a disordered (confined) phase for smaller β values, and an ordered (deconfined) phase for larger β values. The two phases are separated by a first-order phase transition. This is the case for vanishing and small chemical potential. With increasing chemical potential, the transition weakens and turns into a crossover at a critical endpoint. For larger h, there is a crossover only.

We will also consider two closely related models which have a real action, namely the model with imaginary chemical potential, $\mu = i\mu_{\mathrm{I}}$, and the phase-quenched model, obtained by discarding the imaginary part of the action, such that

$$S_F^{\mathrm{pq}} = -h \cosh \mu \sum_x \left(\mathrm{Tr}\, U_x + \mathrm{Tr}\, U_x^\dagger \right). \tag{6.5}$$

In contrast to QCD, the SU(3) spin model does not have a Silver Blaze problem. The Silver Blaze problem [74] refers to the region in the phase diagram where the chemical potential is nonzero but bulk thermodynamic observables, such as the pressure and the density, are μ-independent. This μ-independence requires a precise cancellation which can be highly non-trivial in a numerical approach, as can be seen from studies of the eigenvalues of the Dirac operator [74, 85]. We note here that complex Langevin dynamics has been shown to be able to solve the Silver Blaze problem, in the relativistic Bose gas [55, 56] and in one-dimensional QCD [82]. To see that the Silver

Blaze region is absent in this model, consider the density,

$$\langle n \rangle = \frac{1}{\Omega} \frac{\partial \ln Z}{\partial \mu} = \left\langle h e^{\mu} \mathrm{Tr}\, U_x - h e^{-\mu} \mathrm{Tr}\, U_x^\dagger \right\rangle . \tag{6.6}$$

A nonzero density induces a difference between $\langle \mathrm{Tr}\, U_x \rangle$ and $\langle \mathrm{Tr}\, U_x^\dagger \rangle$. On the other hand, in the Silver Blaze region, $\langle n \rangle = 0$ and $\langle \mathrm{Tr}\, U_x \rangle = \langle \mathrm{Tr}\, U_x^\dagger \rangle$. It is clear from the expression above that it is not possible to simultaneously satisfy these conditions when $\mu \neq 0$, hence the Silver Blaze region is absent.[1] Similarly, the density in the phase-quenched theory,

$$\langle n \rangle_{\mathrm{pq}} = h \sinh \mu \left\langle \mathrm{Tr}\, U_x + \mathrm{Tr}\, U_x^\dagger \right\rangle_{\mathrm{pq}} , \tag{6.7}$$

is nonzero as soon as $\mu > 0$.

The severeness of the sign problem is conventionally estimated via the expectation value of the complex phase factor $e^{i\varphi} = e^{-S}/|e^{-S}|$ in the phase quenched theory,

$$\langle e^{i\varphi} \rangle_{\mathrm{pq}} = \frac{Z}{Z_{\mathrm{pq}}} = e^{-\Omega(f - f_{\mathrm{pq}})}. \tag{6.8}$$

The full and the phase-quenched theory differ as soon as μ is nonzero, which can be seen by performing a Taylor series expansion around $\mu = 0$. To second order in μ, the free energy densities read

$$f(\mu) = f(0) - (c_1 + c_2 h)\, h\mu^2 + \mathcal{O}(\mu^4), \tag{6.9}$$

$$f_{\mathrm{pq}}(\mu) = f(0) - c_1 h \mu^2 + \mathcal{O}(\mu^4), \tag{6.10}$$

[1] For completeness, we recall that $\langle \mathrm{Tr}\, U_x \rangle$ and $\langle \mathrm{Tr}\, U_x^\dagger \rangle$ are both real in the full theory, while at imaginary μ the real parts are equal and the imaginary parts are opposite. In the phase-quenched theory, they are real and identical. Note also that the fermion contribution breaks the Z(3) symmetry of the bosonic sector, hence $\langle \mathrm{Tr}\, U_x \rangle$ and $\langle \mathrm{Tr}\, U_x^\dagger \rangle$ are never strictly zero.

with

$$c_1 = \frac{1}{\Omega} \sum_x \langle \mathrm{Tr}\, U_x \rangle_{\mu=0}\,, \tag{6.11}$$

$$c_2 = \frac{1}{2\Omega} \sum_{xy} \left\langle \mathrm{Tr}\left(U_x - U_x^\dagger\right) \mathrm{Tr}\left(U_y - U_y^\dagger\right) \right\rangle_{\mu=0}. \tag{6.12}$$

Since c_2 is negative [$\mathrm{Tr}\left(U_x - U_x^\dagger\right)$ is imaginary], $f - f_{\mathrm{pq}} \geq 0$, as it should be. Similarly, $\langle n \rangle \leq \langle n \rangle_{\mathrm{pq}}$.

6.3 Discretized complex Langevin dynamics

In order to solve the complex Langevin evolution numerically, we follow Ref. [78] and diagonalize the SU(3) matrices in terms of the angles $\phi_{1,2}$, such that

$$\mathrm{Tr}\, U_x = e^{i\phi_{1x}} + e^{i\phi_{2x}} + e^{-i(\phi_{1x}+\phi_{2x})}, \tag{6.13}$$

$$\mathrm{Tr}\, U_x^\dagger = e^{-i\phi_{1x}} + e^{-i\phi_{2x}} + e^{i(\phi_{1x}+\phi_{2x})}. \tag{6.14}$$

We then have to include the reduced Haar measure and consider the partition function

$$Z = \prod_x \int_{-\pi}^{\pi} d\phi_{1x}\, d\phi_{2x}\, e^{-S_{\mathrm{eff}}}, \tag{6.15}$$

where

$$S_{\mathrm{eff}} = S_B + S_F + S_H, \tag{6.16}$$

with

$$S_H = -\sum_x \ln\left[\sin^2\left(\frac{\phi_{1x} - \phi_{2x}}{2}\right) \sin^2\left(\frac{2\phi_{1x} + \phi_{2x}}{2}\right) \sin^2\left(\frac{\phi_{1x} + 2\phi_{2x}}{2}\right)\right]. \tag{6.17}$$

We note that it is also possible to implement complex Langevin dynamics directly for the SU(3) matrices, see e.g. Refs. [54, 79].

Langevin dynamics provides a stochastic update for the angles ϕ_{ax} ($a =$

1, 2), according to

$$\frac{\partial}{\partial\vartheta}\phi_{ax} = K_{ax} + \eta_{ax}, \qquad K_{ax} = -\frac{\partial S_{\text{eff}}}{\partial\phi_{ax}}, \tag{6.18}$$

where ϑ denotes the Langevin time, K_{ax} is the drift term, and the noise satisfies

$$\langle\eta_{ax}\rangle = 0, \qquad \langle\eta_{ax}\eta_{a'x'}\rangle = 2\delta_{aa'}\delta_{xx'}. \tag{6.19}$$

When the action and hence the drift terms are complex, the angles do not remain real under the Langevin evolution. We therefore write $\phi_{ax} = \phi^{\text{R}}_{ax} + i\phi^{\text{I}}_{ax}$, and consider the following complex Langevin equations, using real noise,

$$\frac{\partial}{\partial\vartheta}\phi^{\text{R}}_{ax} = K^{\text{R}}_{ax} + \eta_{ax}, \qquad K^{\text{R}}_{ax} = -\text{Re}\,\frac{\partial S_{\text{eff}}}{\partial\phi_{ax}}\Big|_{\phi_{ax}\to\phi^{\text{R}}_{ax}+i\phi^{\text{I}}_{ax}}, \tag{6.20}$$

$$\frac{\partial}{\partial\vartheta}\phi^{\text{I}}_{ax} = K^{\text{I}}_{ax}, \qquad K^{\text{I}}_{ax} = -\text{Im}\,\frac{\partial S_{\text{eff}}}{\partial\phi_{ax}}\Big|_{\phi_{ax}\to\phi^{\text{R}}_{ax}+i\phi^{\text{I}}_{ax}}. \tag{6.21}$$

After complexification, we write U_x^{-1} instead of $U_x^\dagger$ in the remainder.

To solve these equations numerically, Langevin time is discretized as $\vartheta = \epsilon n$, where ϵ is the Langevin time step. The standard algorithm discretizing Eq. (6.18) reads[2]

$$\phi_{ax}(n+1) = \phi_{ax}(n) + \epsilon K[\phi_{ax}(n)] + \sqrt{\epsilon}\,\eta_{ax}(n), \tag{6.22}$$

where

$$\langle\eta_{ax}(n)\rangle = 0, \qquad \langle\eta_{ax}(n)\eta_{a'x'}(n')\rangle = 2\delta_{aa'}\delta_{xx'}\delta_{nn'}. \tag{6.23}$$

The contribution to the drift term from the Haar measure requires careful integration. For this we use the adaptive stepsize algorithm of Chapter 2 and Ref. [?].

It is well-known that Langevin dynamics has finite stepsize corrections,

[2]Complexification is obvious and we do not give the discretized equations explicitly.

which are linear in ϵ in the lowest-order discretization given above [50]. It is therefore necessary to extrapolate to zero stepsize. In our previous work [?, 54–56, 80–83], we have only considered the lowest-order algorithm. However, motivated by the results to be presented below, we implemented a higher-order algorithm to improve the stepsize dependence. A standard Runge-Kutta scheme, where the drift terms are improved but the noise is kept as above, will not remove the leading stepsize correction [86]. Instead, it is necessary to modify the noise terms as well. We use the algorithm proposed in Ref. [84] for real Langevin dynamics, which is explicit and easy to implement.[3] It takes the following form

$$\begin{aligned}
\psi_{ax}(n) &= \phi_{ax}(n) + \frac{1}{2}\epsilon K[\phi_{ax}(n)], \\
\tilde{\psi}_{ax}(n) &= \phi_{ax}(n) + \frac{1}{2}\epsilon K[\phi_{ax}(n)] + \frac{3}{2}\sqrt{\epsilon}\,\tilde{\alpha}_{ax}(n), \\
\phi_{ax}(n+1) &= \phi_{ax}(n) + \frac{1}{3}\epsilon\left(K[\psi_{ax}(n)] + 2K[\tilde{\psi}_{ax}(n)]\right) + \sqrt{\epsilon}\,\alpha_{ax}(n).
\end{aligned} \tag{6.24}$$

Here $\tilde{\alpha}_{ax}(n)$ is a random variable taken according to

$$\tilde{\alpha}_{ax}(n) = \frac{1}{2}\alpha_{ax}(n) + \frac{\sqrt{3}}{6}\xi_{ax}(n), \tag{6.25}$$

while $\alpha_{ax}(n)$ and $\xi_{ax}(n)$ are independent Gaussian random variables with variance 2 and vanishing mean, i.e.,

$$\begin{aligned}
\langle \alpha_{ax}(n)\alpha_{a'x'}(n')\rangle &= \langle \xi_{ax}(n)\xi_{a'x'}(n')\rangle = 2\delta_{aa'}\delta_{xx'}\delta_{nn'}, \\
\langle \alpha_{ax}(n)\xi_{a'x'}(n')\rangle &= \langle \alpha_{ax}(n)\rangle = \langle \xi_{ax}(n)\rangle = 0.
\end{aligned} \tag{6.26}$$

In Ref. [84] it was shown analytically, for the case of a real drift term, that with this update the remaining correction is $\mathcal{O}(\epsilon^2)$ for a system with one degree of freedom and $\mathcal{O}(\epsilon^{3/2})$ for a coupled system. In Appendix A we discuss this algorithm in some more detail.

[3] For other approaches, see e.g. Refs. [87, 88].

6.4 Justification and criteria for correctness

In the case of a real action, it can be shown that stochastic quantization and Langevin dynamics is equivalent to standard path integral quantization [50]. As is well-known, such a general statement is lacking in the case of a complex action [50]. Indeed, it can occur that under complex Langevin evolution expectation values converge to a wrong result [44, 46, 47, 52, 53, 81, 89]. It is therefore important to be able to judge the outcome of a complex Langevin process using assessments which are general and can be used in a variety of theories, especially when there are no known results to compare with.

The first assessment employs analyticity at small μ^2: observables, which are even under charge conjugation, should be analytic as a function of μ^2 (in a finite volume) [6, 90]. Results at positive μ^2 can be compared with those where $\mu^2 < 0$, i.e. at imaginary potential, obtained using real Langevin dynamics or any other standard approach. This test is limited to small chemical potentials.

A more formal justification of the complexified dynamics can be found in Refs. [80, 83]. Here we summarize that discussion briefly in order to arrive at the criteria for correctness developed in Ref. [83]. We consider expectation values with respect to the real and positive probability distribution $P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}; \vartheta]$, sampled by the stochastic process,

$$\langle O \rangle_{P(\vartheta)} = \frac{\int D\phi^{\mathrm{R}} D\phi^{\mathrm{I}}\, P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}; \vartheta] O[\phi^{\mathrm{R}} + i\phi^{\mathrm{I}}]}{\int D\phi^{\mathrm{R}} \phi^{\mathrm{I}}\, P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}; \vartheta]}. \tag{6.27}$$

Here ϑ is the Langevin time. With the help of the Langevin equations for ϕ^{R} and ϕ^{I}, one finds the Fokker-Planck equation for $P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}; \vartheta]$,

$$\frac{\partial P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}; \vartheta]}{\partial \vartheta} = L^T P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}}; \vartheta], \qquad L^T = \frac{\partial}{\partial \phi^{\mathrm{R}}} \left[\frac{\partial}{\partial \phi^{\mathrm{R}}} - K^{\mathrm{R}} \right] - \frac{\partial}{\partial \phi^{\mathrm{I}}} K^{\mathrm{I}}. \tag{6.28}$$

We also consider expectation values with respect to a complex weight $\rho[\phi;\vartheta]$,

$$\langle O\rangle_{\rho(\vartheta)} = \frac{\int D\phi\, \rho[\phi;\vartheta] O[\phi]}{\int D\phi\, \rho[\phi;\vartheta]}, \tag{6.29}$$

where ϕ is real and $\rho[\phi;\vartheta]$ satisfies

$$\frac{\partial \rho[\phi;\vartheta]}{\partial \vartheta} = L_0^T \rho[\phi;\vartheta], \qquad L_0^T = \frac{\partial}{\partial \phi}\left[\frac{\partial}{\partial \phi} + \frac{\partial S}{\partial \phi}\right]. \tag{6.30}$$

Note that the Fokker-Planck operators L^T, acting on the real density $P[\phi^{\mathrm{R}}, \phi^{\mathrm{I}};\vartheta]$, and L_0^T, acting on the complex density $\rho[\phi;\vartheta]$, should be distinguished. Furthermore, Eq. (6.30) has a stationary solution, $\rho[\phi] \sim \exp(-S)$, whereas for Eq. (6.28) no generic stationary solution is known.

Employing that the only permissible observables are holomorphic and making use of partial integration, one can show that expectation values with respect to the two densities are equal,

$$\langle O\rangle_{P(\vartheta)} = \langle O\rangle_{\rho(\vartheta)}. \tag{6.31}$$

If subsequently one can show that $\rho[\phi;\vartheta]$ reaches the unique stationary solution $\sim \exp(-S)$ in the limit that $\vartheta \to \infty$, the use of complex Langevin dynamics is justified [80].

The equivalence in Eq. (6.31) relies on the ability to do partial integration without receiving contributions from boundary terms, i.e. the distributions should be well localized and decay strongly. It was shown in Ref. [83] that this condition can be expressed as a set of criteria on holomorphic observables O, which take the form

$$\langle \tilde{L} O\rangle = 0. \tag{6.32}$$

Here $\tilde{L}$ denotes the Langevin operator

$$\tilde{L} = \left(\frac{\partial}{\partial \phi} + K\right)\frac{\partial}{\partial \phi}, \qquad K = -\frac{\partial S}{\partial \phi}, \tag{6.33}$$

which depends on holomorphic degrees of freedom ϕ. Although it differs from L and L_0, the action of $\tilde{L}$ on holomorphic observables agrees with that of L. The expectation value in Eq. (6.32) is taken with respect to the weight P in the limit that the Langevin process has equilibrated ($\vartheta \to \infty$). In principle, the criteria (6.32) should be satisfied for a large enough set of holomorphic observables [83].

Adapting this to the model at hand, $\tilde{L}$ reads

$$\tilde{L} = \sum_{x,a} \left(\frac{\partial}{\partial \phi_{ax}} + K_{ax} \right) \frac{\partial}{\partial \phi_{ax}}. \tag{6.34}$$

We will consider only local observables and denote these as $O[\phi_{1x}, \phi_{2x}] = O_x$. We can then write

$$\tilde{L}O_x = \sum_a \left(O_x^{a''} + K_{ax} O_x^{a'} \right), \tag{6.35}$$

where

$$O_x^{a'} = \frac{\partial O_x}{\partial \phi_{ax}}, \qquad O_x^{a''} = \frac{\partial^2 O_x}{\partial \phi_{ax}^2}. \tag{6.36}$$

In terms of the real and imaginary parts, this yields explicitly

$$\mathrm{Re}\, \tilde{L}O_x = \sum_a \left(\mathrm{Re}\, O_x^{a''} + K_{ax}^{\mathrm{R}} \mathrm{Re}\, O_x^{a'} - K_{ax}^{\mathrm{I}} \mathrm{Im}\, O_x^{a'} \right), \tag{6.37}$$

$$\mathrm{Im}\, \tilde{L}O_x = \sum_a \left(\mathrm{Im}\, O_x^{a''} + K_{ax}^{\mathrm{R}} \mathrm{Im}\, O_x^{a'} + K_{ax}^{\mathrm{I}} \mathrm{Re}\, O_x^{a'} \right), \tag{6.38}$$

and the criteria read

$$\langle \mathrm{Re}\, \tilde{L}O_x \rangle = 0, \qquad \langle \mathrm{Im}\, \tilde{L}O_x \rangle = 0. \tag{6.39}$$

6.5 Results and justification

In this section we present a number of results obtained with complex Langevin dynamics. As mentioned earlier, our goal is not deliver a detailed study of

critical properties and the phase structure; rather the aim of this study is to assess the reliability of the complex Langevin algorithm, using the criteria discussed in the previous section.

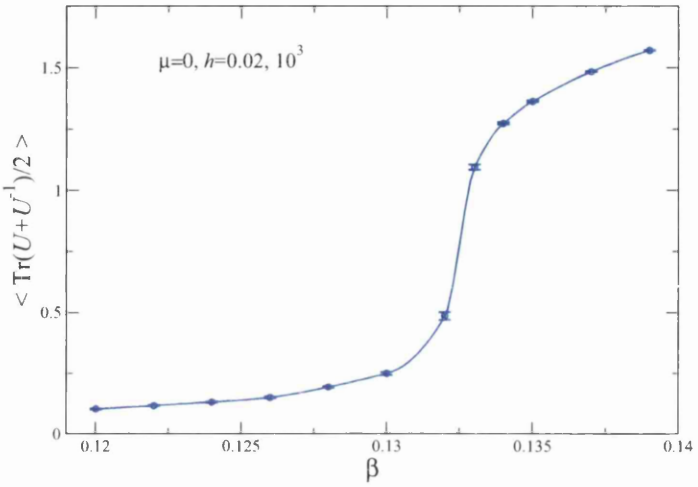


Figure 6.1: $\langle \mathrm{Tr}\ (U_x + U_x^{-1})/2 \rangle$ as a function of β at $\mu = 0$ and $h = 0.02$ on a 10^3 lattice, using real Langevin dynamics.

The results we show here are obtained using a relatively small value for the fermion coupling, $h = 0.02$, so that there is a clear transition between the ordered and the disordered phase. We consider β values between 0.12 and 0.139; the critical β value for the fermion coupling we use is around 0.1324 at $\mu = 0$. This is demonstrated in Fig. 6.1, where we show $\langle \mathrm{Tr}\ (U_x + U_x^{-1})/2 \rangle = \langle \mathrm{Tr}\, U_x \rangle$ as a function of β at $\mu = 0$ on a 10^3 lattice.

As a first test, we probe the transition by varying μ instead of β. As mentioned above, observables which are invariant under charge conjugation should, in a finite volume, be analytic in μ^2 [6,90]. This yields the possibility to compare results at positive μ^2 with those at negative μ^2, corresponding to imaginary potential. Since in this case the action is real, real Langevin dynamics can be used, which is theoretically well founded. In Fig. 6.2 we show $\langle \mathrm{Tr}\ (U_x + U_x^{-1})/2 \rangle$ as a function of μ^2 for eight different β values. We

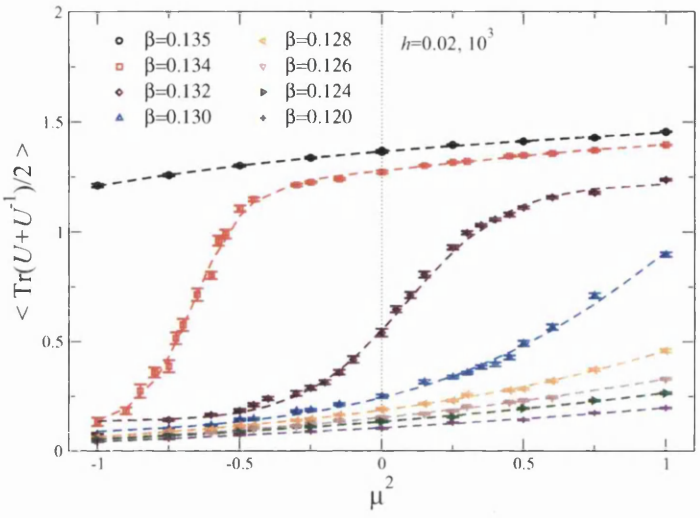


Figure 6.2: Analyticity in μ^2: $\langle \mathrm{Tr}\,(U_x + U_x^{-1})/2 \rangle$ as a function of μ^2 for various β values with $h = 0.02$ on a 10^3 lattice. Data at imaginary μ (with $\mu^2 \leq 0$) has been obtained with real Langevin dynamics, data at real μ (with $\mu^2 > 0$) with complex Langevin dynamics.

observe smooth behaviour as μ^2 is increased. This is an indication that complex Langevin dynamics works well. We note that this is true in both phases as well as in the transition region. This is in contrast to the case of the XY model recently studied using complex Langevin dynamics, where correct results were obtained in only part of the phase diagram [81].

The strength of the transition weakens as μ^2 increases and, vice versa, increases as μ^2 decreases [38, 78, 79]. This can be seen in Fig. 6.3, where we show the Langevin time evolution of $\langle \mathrm{Tr}\,(U_x + U_x^{-1})/2 \rangle$ at $\mu^2 = -0.65$ and $\beta = 0.134$ (left) and $\mu^2 = 0.1$ and $\beta = 0.132$ (right). We observe clear first order behaviour at $\mu^2 = -0.65$, while at $\mu^2 = 0.1$ the transition is much weaker.

Next we consider the density as a function of μ in the full and the phase-quenched theory. In Fig. 6.4 we show the density for chemical potentials up to $\mu = 3.5$, at $\beta = 0.125$. For this β value the model is in the disordered phase

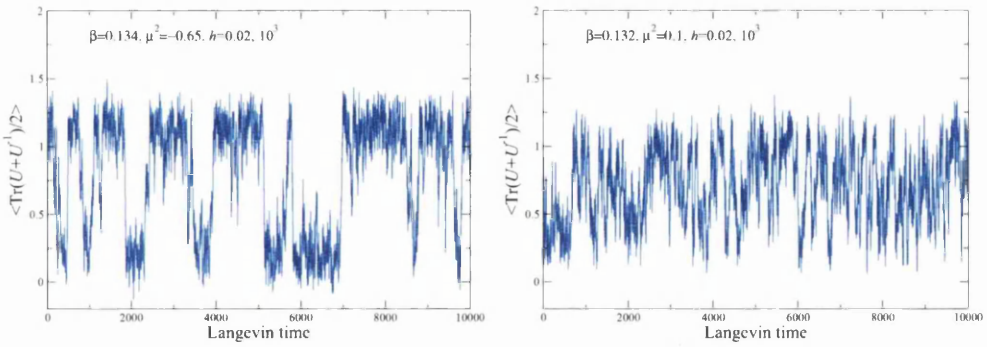


Figure 6.3: Langevin time evolution of $\langle \mathrm{Tr}\,(U_x + U_x^{-1})/2 \rangle$ in the transition region, at imaginary chemical, $\mu^2 = -0.65$ and $\beta = 0.134$ (left) and real chemical potential, $\mu^2 = 0.1$ and $\beta = 0.132$ (right). The other parameters are as above.

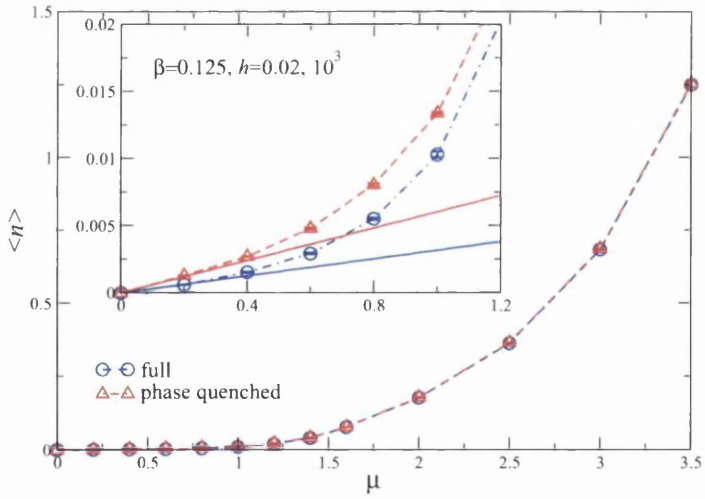


Figure 6.4: Density $\langle n \rangle$ in the full and the phase-quenched theory as a function of μ at $\beta = 0.125$ and $h = 0.02$ on a 10^3 lattice. The inset shows a close-up of the small μ region. The lines are the predicted linear dependence for small μ, evaluated at $\mu = 0$.

at smaller μ values and in the ordered phase at larger μ values. The densities in the full and phase-quenched theories are similar, but not equal. We recall that there is no Silver Blaze region in this model. This can be seen in the inset, which shows a close-up: the density in the full theory is below the one in the phase-quenched theory, but it is nonzero (we have verified that there are no visible finite-size effects remaining). The lines indicate the expected linear dependence of the densities on μ, using the lowest-order Taylor series expansion, see Eqs. (6.9, 6.10),

$$\langle n \rangle = 2\,(c_1 + c_2 h)\,h\mu + \mathcal{O}(\mu^3), \tag{6.40}$$

where the coefficients $c_{1,2}$ have been defined in Eqs. (6.11, 6.12). In the phase-quenched theory the term with c_2 is absent. We have computed the coefficients and find

$$c_1 = 0.1446(21), \qquad c_2 = -3.534(72). \tag{6.41}$$

Using these coefficients yields the straight lines in the inset of Fig. 6.4, justifying the results of complex Langevin dynamics.

In Fig. 6.5 we show $\langle \mathrm{Tr}\, U \rangle$ and $\langle \mathrm{Tr}\, U^{-1} \rangle$ as a function of μ in the full theory, using the same parameters as in Fig. 6.4. Recall that $\langle \mathrm{Tr}\, U \rangle$ and $\langle \mathrm{Tr}\, U^{-1} \rangle$ are both real and that one expects $\langle \mathrm{Tr}\, U \rangle < \langle \mathrm{Tr}\, U^{-1} \rangle$, due to the nonzero density. At small μ, the linear dependence on μ can again be expressed in terms of the coefficients $c_{1,2}$ and we find

$$\langle \mathrm{Tr}\, U \rangle = c_1 + c_2 h\mu + \mathcal{O}(\mu^2), \tag{6.42}$$

$$\langle \mathrm{Tr}\, U^{-1} \rangle = c_1 - c_2 h\mu + \mathcal{O}(\mu^2). \tag{6.43}$$

This yields the straight lines in the inset of Fig. 6.5, justifying again the results of complex Langevin dynamics. In the phase-quenched theory, $\langle \mathrm{Tr}\, U \rangle_{\mathrm{pq}}$ and $\langle \mathrm{Tr}\, U^{-1} \rangle_{\mathrm{pq}}$ are equal and slightly below $\langle \mathrm{Tr}\,(U + U^{-1})/2 \rangle$ in the full

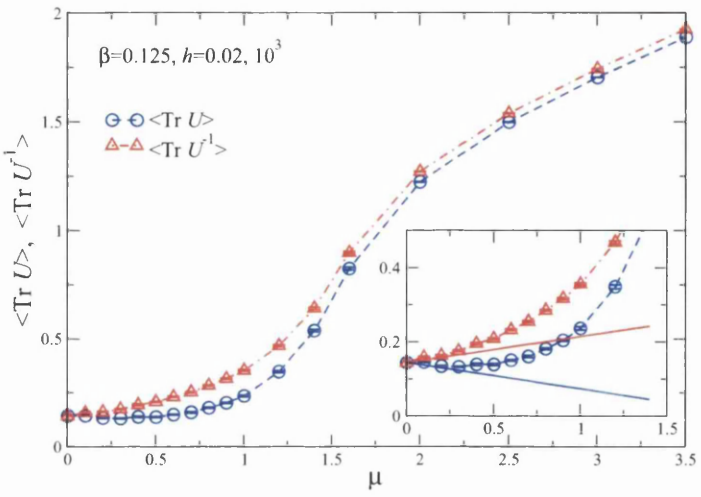


Figure 6.5: $\langle \mathrm{Tr}\, U_x \rangle$ and $\langle \mathrm{Tr}\, U_x^{-1} \rangle$ as a function of μ in the full theory. The parameters are as in Fig. 6.4. The inset shows a close-up of the small μ region. The lines are the predicted linear dependence for small μ, evaluated at $\mu = 0$.

theory (not shown).

The average phase factor, indicating the severeness of the sign problem, is shown in Fig. 6.6 for a typical choice of parameters. The lines indicate the behaviour expected at small chemical potential,

$$\langle e^{i\varphi} \rangle_{\mathrm{pq}} = e^{-\Omega \Delta f}, \qquad \Delta f = f - f_{\mathrm{pq}} = -c_2 h^2 \mu^2 + \mathcal{O}(\mu^4). \tag{6.44}$$

As in preceding studies [54–56, 81, 82], we have not observed a correlation between the severeness of the sign problem and the efficiency of the complex Langevin algorithm. We also note that the average phase factor behaves in a non-monotonic manner as a function of μ in the transition region.

In order to assess complex Langevin dynamics in detail for larger μ values, we now focus on two points in the phase diagram: $\beta = 0.125$, $\mu = 1$ in the disordered phase and $\beta = 0.125$, $\mu = 3$ in the ordered phase. To

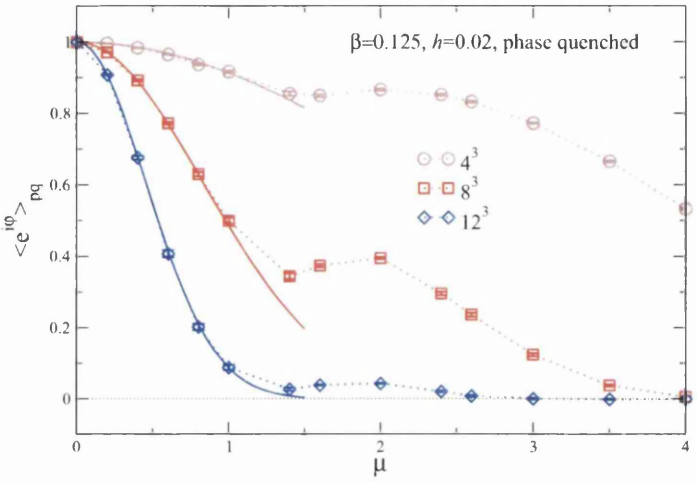


Figure 6.6: Average phase factor in the phase-quenched theory $\langle e^{i\varphi} \rangle_{\rm pq}$ as a function of μ, for various volumes at $\beta = 0.125$ and $h = 0.02$. The lines indicate the expected behaviour using the leading μ^2 term at small μ.

control the statistical error we have carried out simulations using a total Langevin trajectory of length 10,000 (after discarding the thermalization stage) in the disordered phase; in the ordered phase fluctuations are smaller and a Langevin trajectory of 5,000 is sufficient. Errors are determined with a jackknife analysis. We have used a number of stepsizes, from $\epsilon = 0.001$ down to $\epsilon = 0.00005$, employing both the standard lowest-order algorithm and the improved higher-order algorithm. The results are collected in Tables 1 and 2 in Appendix B.

In Fig. 6.7 we show $\langle \mathrm{Tr}\, U \rangle$ and $\langle \mathrm{Tr}\, U^{-1} \rangle$ as a function of the Langevin stepsize for $\mu = 1$ (left) and 3 (right). Statistical fluctuations in the disordered phase are larger, even though the Langevin trajectory is twice as long. For the lowest-order algorithm stepsize dependence is clearly visible, as expected. The dotted lines indicate a linear fit using the data at the four smallest stepsizes. In the case of the higher-order algorithm, there appears to be no stepsize dependence visible; the dashed lines indicate the average

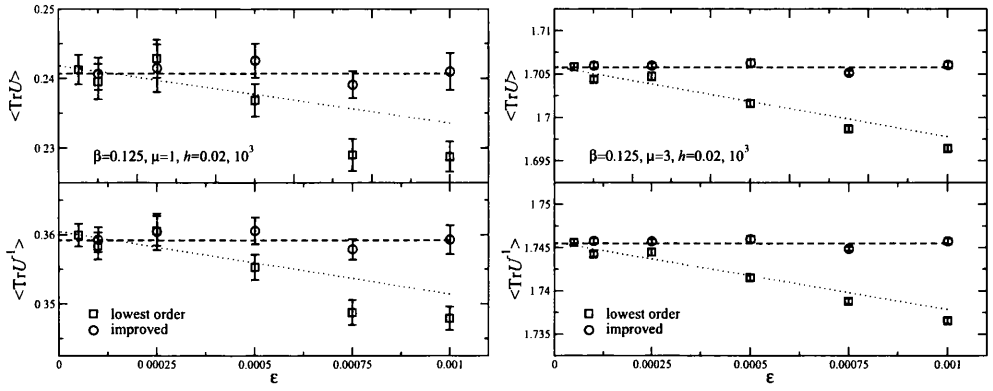


Figure 6.7: Stepsize dependence of $\langle \mathrm{Tr}\, U_x \rangle$ (top panes) and $\langle \mathrm{Tr}\, U_x^{-1} \rangle$ (bottom panes) at $\mu = 1$ (left) and 3 (right) on a 10^3 lattice for $\beta = 0.125$ and $h = 0.02$, using both the standard lowest-order and the improved algorithm.

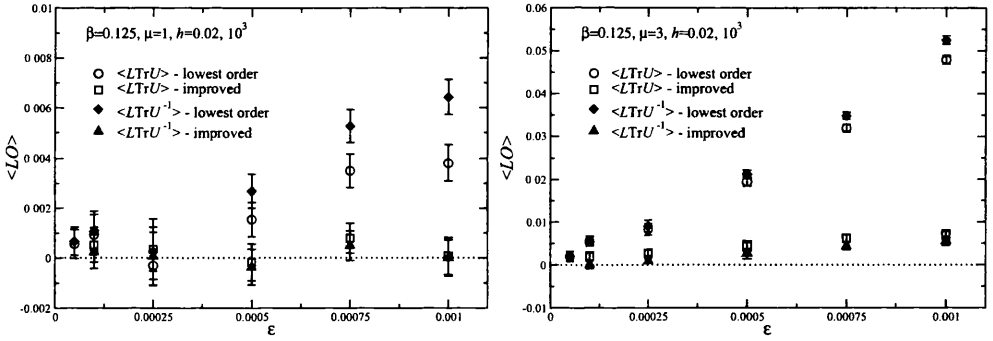


Figure 6.8: Stepsize dependence of the real part of $\langle L \mathrm{Tr}\, U_x \rangle$ and $\langle L \mathrm{Tr}\, U_x^{-1} \rangle$ at $\mu = 1$ (left) and 3 (right), using both the standard and the improved algorithm. Other parameters as in Fig. 6.7.

of the five data points in each case.[4] Importantly, we note that the results from both algorithms are consistent in the limit $\epsilon \to 0$, see also Appendix B.

In order to justify these results, we have computed $\langle \tilde{L} O \rangle$ (which should be equal to zero), where $O = \mathrm{Tr}\, U, \mathrm{Tr}\, U^{-1}$ and n. Since these observables

[4]Theoretically, corrections of $\mathcal{O}(\epsilon^{3/2})$ are expected [84].

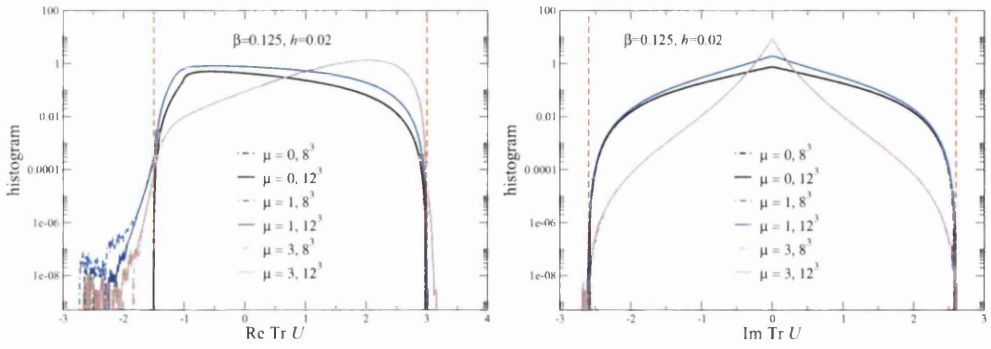


Figure 6.9: Histograms for $\mathrm{Re}\,\mathrm{Tr}\,U_x$ (left) and $\mathrm{Im}\,\mathrm{Tr}\,U_x$ (right) at $\beta = 0.125$, $h = 0.02$ and $\mu = 0, 1, 3$ on 8^3 and 12^3. The vertical lines denote the boundaries in SU(3), i.e. without complexification. Note the vertical logarithmic scale.

are holomorphic, we drop the tilde on the L from now on. Note that $\langle Ln \rangle$ is not independent, since n is a linear combination of $\mathrm{Tr}\,U$ and $\mathrm{Tr}\,U^{-1}$. The imaginary parts of $\langle LO \rangle$ are consistent with zero. The stepsize dependence of the real parts is shown in Fig. 6.8 for $\mu = 1$ (left) and 3 (right). Note the different vertical scale: the stepsize dependence is stronger in the ordered phase.[5] For the lowest-order algorithm there are again clear finite-stepsize corrections, which vanish in the limit that $\epsilon \to 0$. In the case of the higher-order algorithm, finite-stepsize corrections are much smaller or even absent. We find that $\langle LO \rangle$ goes to zero in the limit that the stepsize is taken to zero. This observation is a necessary requirement for the applicability of complex Langevin dynamics. Interestingly, larger finite-stepsize corrections correspond to larger deviations of $\langle LO \rangle$ from zero. It turns out that this is also seen when using real Langevin dynamics, e.g. in the phase-quenched theory. We conclude therefore that computations of $\langle LO \rangle$ yield a sensitive test to quantify finite-stepsize errors.

As a final assessment, we discuss the extent to which the complexified field

[5]Larger stepsize corrections in the ordered phase at larger values of μ are also seen in the Bose gas, where the stepsize is effectively enhanced as $e^{\mu}\epsilon$ [56].

space is explored. A sufficiently localized distribution in the imaginary field direction is required for the formal justification to hold [80,83]. In Fig. 6.9 we show histograms for $\mathrm{Re}\,\mathrm{Tr}\,U$ (left) and $\mathrm{Im}\,\mathrm{Tr}\,U$ (right), obtained by binning the data sampled during the Langevin process, for three different μ values and two lattice volumes at $\beta = 0.125$. For real dynamics, i.e. when the angles $\phi_{1,2}$ are real and $U \in \mathrm{SU}(3)$, $\mathrm{Tr}\,U$ is complex-valued, taking values in a triangular shape with corners at $3e^{2q\pi i/3}$ $(q = 0, 1, 2)$. The corresponding boundaries are shown in the figures as vertical dashed lines and the histograms at $\mu = 0$ are contained within these boundaries. After complexification, $\phi_{1,2}$ are no longer proper angles and the dynamics takes place in the larger group $\mathrm{SL}(3,\mathbb{C})$. At nonzero μ, we observe that the SU(3) boundaries are indeed crossed, as required, but that the distribution appears to remain localized (note the vertical logarithmic scale). The tails of the distributions are noisy, as they are visited during the Langevin process very rarely. There is very little volume dependence. We also note that the histograms at $\mu = 1$ (in the disordered phase) resemble the histograms at $\mu = 0$, while at $\mu = 3$ (in the ordered phase) they are significantly different, which is reflected in the larger expectation value of $\langle \mathrm{Tr}\,U \rangle$. The histograms for $\mathrm{Im}\,\mathrm{Tr}\,U$ are symmetric within numerical uncertainty, since $\langle \mathrm{Tr}\,U \rangle$ is real.

Finally, in Fig. 6.10 we show the histogram for $\phi^{\mathrm{I}} = \{\phi_1^{\mathrm{I}}, \phi_2^{\mathrm{I}}\}$. For real Langevin dynamics at $\mu = 0$, $\phi^{\mathrm{I}} \equiv 0$. At finite μ, nonzero values are generated by the complex drift term. We observe that the distribution drops exponentially, over many decades, before the signal becomes noisy. The straight dashed lines indicate $P(\phi^{\mathrm{I}}) \sim e^{-b|\phi^{\mathrm{I}}|}$ with $b = 35, 45$, and are meant to guide the eye. Note that the exponential drop is considerably faster than in the U(1) model studied in Refs. [80,83], where $b \sim 2$ and complex Langevin dynamics failed. This fast drop and localization of the distribution is another requirement for the applicability of complex Langevin dynamics.[6]

[6] An open question is what happens to expectation values of the form $\langle \mathrm{Tr}\,U^k \rangle$ with k large. These observables contain terms of the form $e^{-k\phi^{\mathrm{I}}} \cos(k\phi^{\mathrm{R}})$ and the presence of the rapidly oscillating cosine should be taken into consideration, see also Ref. [83].

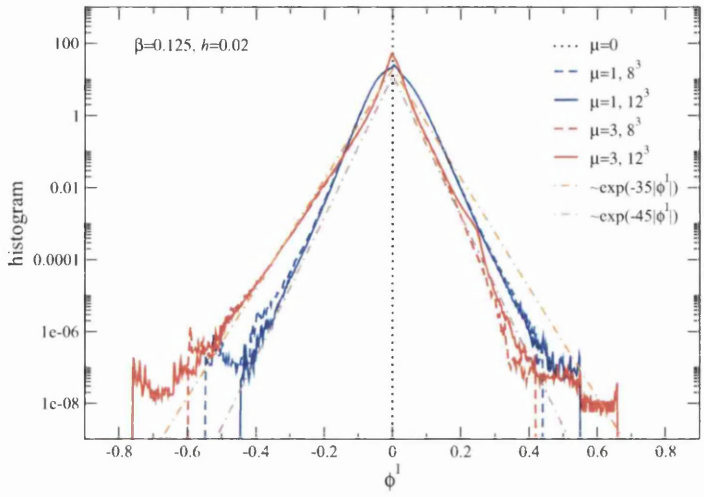


Figure 6.10: Histogram $P(\phi^{\mathrm{I}})$, where $\phi^{\mathrm{I}} = \{\phi_1^{\mathrm{I}}, \phi_2^{\mathrm{I}}\}$, at $\beta = 0.125$, $h = 0.02$ and $\mu = 0, 1, 3$ on 8^3 and 12^3. When $\mu = 0$, $\phi^{\mathrm{I}} \equiv 0$. The dashed straight lines are $P(\phi^{\mathrm{I}}) \sim e^{-b|\phi^{\mathrm{I}}|}$ with $b = 35, 45$. Note the vertical logarithmic scale.

6.6 Effective one link model

Since it is always possible to consider each lattice site coming into equilibrium with a heat bath from the surrounding sites, it is sufficient that if the results from a complex Langevin simulations agree with the exact results for all complex β, then the full model must also converge correctly. This is of course a stronger requirement than is strictly necessary, since not all possible complex β will be sampled by the full model. In order to better understand the convergence of the complex Langevin process, one can write the nearest-neighbour terms in the action to focus on a single site, at position x,

$$S_B = -\beta \sum_x \left[\mathrm{Tr}\, U_x \left(\sum_\nu \mathrm{Tr}\, U^\dagger_{x+\hat{\nu}} + \mathrm{Tr}\, U^\dagger_{x-\hat{\nu}} \right) + \mathrm{Tr}\, U^\dagger_x \left(\sum_\nu \mathrm{Tr}\, U_{x+\hat{\nu}} + \mathrm{Tr}\, U_{x-\hat{\nu}} \right) \right]. \tag{6.45}$$

This motivates a simple toy model, where the coupling β is replaced with a complex parameter which encompasses the effect of the complex neighbours. Writing $\beta_{\text{eff}} e^{i\gamma} = 6\beta u$, with u representing the interactions with the neighbours and is in general a complex number. Then the toy model has the action

$$S = -\beta_1 \mathrm{Tr}\, U - \beta_2 \mathrm{Tr}\, U^\dagger, \tag{6.46}$$

with the parameters given by

$$\beta_1 = \beta_{\text{eff}} e^{i\gamma} + h e^{\mu}, \tag{6.47}$$

$$\beta_2 = \beta_{\text{eff}} e^{-i\gamma} + h e^{-\mu} \tag{6.48}$$

Typical values for u can be computed from simulations of the full lattice model, where histograms show that even when $\mu \neq 0$ the range of values that u can take remain largely within the confines of $SU(3)$. Specifically, this means that $\beta_{\text{eff}} \lesssim 3$ and $-\pi \leq \gamma \leq \pi$.

As before, the elements of SU(3) can be reduced to two angles, which requires introducing the reduced Haar measure, so that the total action has again a contribution from $S_H = -\ln H(\phi_1, \phi_2)$,

$$H(\phi_1, \phi_2) = \sin^2\left(\frac{\phi_1 - \phi_2}{2}\right) \sin^2\left(\frac{2\phi_1 + \phi_2}{2}\right) \sin^2\left(\frac{\phi_1 + 2\phi_2}{2}\right). \tag{6.49}$$

The exact result is given by the double integral,

$$\langle O \rangle = \frac{\iint_{-\pi}^{\pi} O(\phi_1, \phi_2) H(\phi_1, \phi_2) e^{-S(\phi_1, \phi_2)} d\phi_1 d\phi_2}{\iint_{-\pi}^{\pi} H(\phi_1, \phi_2) e^{-S(\phi_1, \phi_2)} d\phi_1 d\phi_2}, \tag{6.50}$$

which can be computed directly by numerical integration.

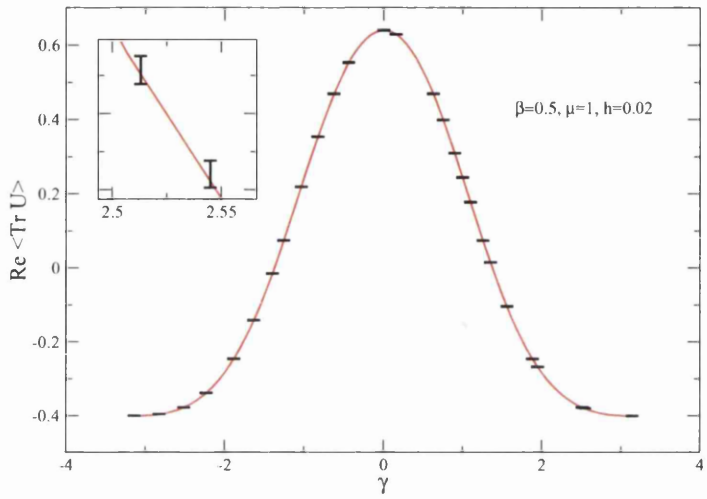
Taking the parameters in the toy model the same as in the full simulation, so that $\mu = 1$ and $h = 0.02$, the complex Langevin results are found to agree with the exact results for all complex β, see Figure 6.6. This satisfies the strongest requirement, that complex Langevin works for all complex β, which

is therefore sufficient for correct convergence of the full model.

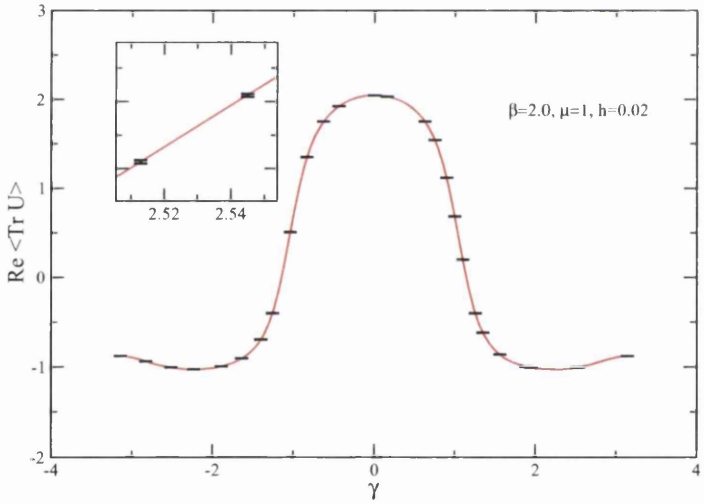
6.7 Summary and outlook

In this chapter we revisited the SU(3) spin model, an effective dimensionally reduced Polyakov loop model for QCD in the strong-coupling and heavy-quark limit, at nonzero chemical potential. To handle the sign problem we employed complex Langevin dynamics, paying special attention to the justification of the method. Using analyticity at small μ^2 (Taylor series expansion and smoothness in μ^2), formal criteria for correctness, and localization of distributions in the complexified space, we arrive at the conclusion that complex Langevin dynamics is reliable in both the ordered and the disordered phase, including the critical region. This should be contrasted with the case of the XY model, where correct results were obtained in the ordered phase but neither in the disordered phase nor in the transition region [81]. In the XY model this failure was detected by an apparent lack of analyticity at small μ^2 and the presence of very broad, slowly decaying distributions (as well as by comparing to results obtained in the world line formalism [91]). We can therefore conclude that the assessments employed here can be used constructively to rule out or support the applicability of complex Langevin dynamics.[7] We emphasize that these tests are generally applicable and not specific to the theory considered here. Besides supporting the results of complex Langevin dynamics, we found that the criteria for correctness are also relevant for real Langevin dynamics, as they show clear sensitivity to finite-stepsize effects. In order to eliminate the leading-order stepsize dependence, we have successfully implemented a simple higher-order algorithm and found it to remove essentially all stepsize dependence in the observables.

[7]Nevertheless, it will still be interesting to compare with results obtained in the flux formulation [37, 38].



(a) $\beta = 0.5$



(b) $\beta = 2.0$

Figure 6.11: Plot of $\mathrm{Re}\,\langle TrU \rangle$ using $\beta_{\mathrm{eff}} = 0.5, 2$ over a range of phases $-\pi \leq \gamma \leq \pi$, solid lines are computed directly from the integral. The inset shows the agreement with the exact result to within small errors.

A Higher-order algorithm

In this appendix, we discuss the higher-order algorithm (6.24) in some more detail. In Ref. [84] the algorithm is constructed by considering a single update step. It was also shown that the set (6.24) is part of a more general update rule. In order to complement the analysis of Ref. [84], we demonstrate the algorithm here using a simple linear kernel. We emphasize that the analysis in Ref. [84] is for general nonlinear drift term. For notational simplicity, we use a single degree of freedom ϕ.

The goal is to solve to higher accuracy

$$\phi_{n+1} = \phi_n + \epsilon K(\phi_n) + \sqrt{\epsilon}\,\eta_n, \qquad \langle \eta_n \eta_{n'} \rangle = 2\delta_{nn'}. \tag{A.1}$$

Consider the linear kernel $K = -\omega\phi$. The solution of the stochastic process, with vanishing initial conditions, is then given by

$$\phi_n = \sqrt{\epsilon} \sum_{i=0}^{n-1} (1 - \epsilon\omega)^{n-1-i}\, \eta_i, \tag{A.2}$$

leading to

$$\lim_{n\to\infty} \langle \phi_n \phi_n \rangle = \frac{1}{\omega} \frac{\langle \eta_i \eta_i \rangle}{2 - \epsilon\omega} = \frac{1}{\omega} \left(1 + \frac{1}{2}\epsilon\omega + \dots \right), \tag{A.3}$$

indicating the linear stepsize dependence (we always assume $\epsilon\omega < 1$).

Ref. [84] proposes the following update

$$\begin{aligned} \psi_n &= \phi_n + \frac{1}{2}\epsilon K(\phi_n) + k\sqrt{\epsilon}\,\tilde{\alpha}_n, \\ \tilde{\psi}_n &= \phi_n + \frac{1}{2}\epsilon K(\phi_n) + l\sqrt{\epsilon}\,\tilde{\alpha}_n, \\ \phi_{n+1} &= \phi_n + \epsilon \left[a K(\psi_n) + b K(\tilde{\psi}_n) \right] + \sqrt{\epsilon}\,\alpha_n, \end{aligned} \tag{A.4}$$

where the coefficients a, b, k and l are to be determined, and the noise satisfies

$$\tilde{\alpha}_n = \frac{1}{2}\alpha_n + \frac{\sqrt{3}}{6}\xi_n, \tag{A.5}$$

and

$$\langle \alpha_n \alpha_{n'} \rangle = \langle \xi_n \xi_{n'} \rangle = 2\delta_{nn'}, \qquad \langle \alpha_n \xi_{n'} \rangle = \langle \alpha_n \rangle = \langle \xi_n \rangle = 0. \tag{A.6}$$

Note that

$$\langle \tilde{\alpha}_n \tilde{\alpha}_{n'} \rangle = \frac{2}{3}\delta_{nn'}, \qquad \langle \alpha_n \tilde{\alpha}_{n'} \rangle = \delta_{nn'}. \tag{A.7}$$

Straightforward substitution in the case of the linear kernel gives

$$\phi_{n+1} = \phi_n - \epsilon\tilde{\omega}\phi_n + \sqrt{\epsilon}\,\tilde{\eta}_n, \tag{A.8}$$

with

$$\tilde{\omega} = \omega(a+b)\left(1 - \frac{1}{2}\epsilon\omega\right), \qquad \tilde{\eta}_n = \alpha_n - (ak+bl)\epsilon\omega\tilde{\alpha}_n. \tag{A.9}$$

Noting that

$$\langle \tilde{\eta}_n \tilde{\eta}_{n'} \rangle = 2\left(1 - (ak+bl)\epsilon\omega + \frac{1}{3}(ak+bl)^2(\epsilon\omega)^2\right)\delta_{nn'}, \tag{A.10}$$

we find, see Eq. (A.3),

$$\lim_{n\to\infty}\langle \phi_n \phi_n \rangle = \frac{1}{\tilde{\omega}}\frac{\langle \tilde{\eta}_i \tilde{\eta}_i \rangle}{2 - \epsilon\tilde{\omega}} = \frac{1}{\omega}\left(\frac{1}{a+b} + \frac{1 + a + b - 2(ak+bl)}{2(a+b)}\epsilon\omega + \ldots\right). \tag{A.11}$$

We can now determine the coefficients and take

$$a + b = 1, \qquad ak + bl = 1. \tag{A.12}$$

The resulting expectation value is

$$\lim_{n\to\infty} \langle \phi_n \phi_n \rangle = \frac{1}{\omega} \left(1 - \frac{\epsilon^2 \omega^2}{6} + \dots \right), \tag{A.13}$$

with a remaining correction of $\mathcal{O}(\epsilon^2)$. The two conditions above do not fully determine the coefficients a, b, k, l. Ref. [84] argues that a further condition

$$ak^2 + bl^2 = \frac{3}{2}, \tag{A.14}$$

follows from minimizing the local truncation error, which is beyond the scope of the linear example discussed here. In the main part of the chapter, the coefficients are taken as

$$a = \frac{1}{3}, \qquad b = \frac{2}{3}, \qquad k = 0, \qquad l = \frac{3}{2}, \tag{A.15}$$

which satisfies the constraints above.

B Tables

In this Appendix we list the results for the stepsize dependence obtained at $\beta = 0.125$ and $h = 0.02$ on a 10^3 lattice, for $\mu = 1$ (Table 1) and $\mu = 3$ (Table 2). The total Langevin time is 10,000 for $\mu = 1$ and 5,000 for $\mu = 3$, after discarding the thermalization stage.

Every table shows the real part of the three observables $\langle \mathrm{Tr}\, U_x \rangle$, $\langle \mathrm{Tr}\, U_x^{-1} \rangle$ and $\langle n \rangle$ in the upper part, and the real part of the criteria $\langle L \mathrm{Tr}\, U_x \rangle$, $\langle L \mathrm{Tr}\, U_x^{-1} \rangle$ and $\langle Ln \rangle$ in the lower part, for both the "lowest-order" and the"improved" algorithm. Here lowest-order algorithm refers to the standard discretization (6.22), which has corrections that are linear in the stepsize; improved algorithm refers to the higher-order algorithm (6.24) of Ref. [84]. In the case of the lowest-order algorithm we performed a linear extrapolation, using the values at the four smallest stepsizes. Since there is very little stepsize depen-

dence left in the case of the improved algorithm, the average is shown.

	ϵ	$\langle \mathrm{Tr}\, U \rangle$	$\langle \mathrm{Tr}\, U^{-1} \rangle$	$\langle n \rangle$
	0.001	0.2288(22)	0.3479(17)	0.00988(11)
lowest	0.00075	0.2290(23)	0.3488(18)	0.00988(11)
order	0.0005	0.2369(23)	0.3553(18)	0.01026(11)
algorithm	0.00025	0.2429(27)	0.3606(21)	0.01055(13)
	0.0001	0.2396(26)	0.3584(20)	0.01039(12)
	0.00005	0.2413(21)	0.3600(16)	0.01047(10)
	extrapolation	0.2419(19)	0.3605(15)	0.010497(92)
	0.001	0.2410(27)	0.3593(21)	0.01046(13)
improved	0.00075	0.2391(20)	0.3579(15)	0.01037(10)
algorithm	0.0005	0.2426(25)	0.3606(19)	0.01053(12)
	0.00025	0.2415(34)	0.3604(26)	0.01048(17)
	0.0001	0.2407(23)	0.3593(18)	0.01044(11)
	average	0.2407(11)	0.35921(84)	0.010443(53)
	ϵ	$\langle L\mathrm{Tr}\, U \rangle$	$\langle L\mathrm{Tr}\, U^{-1} \rangle$	$\langle Ln \rangle$
	0.001	0.00381(71)	0.00644(70)	0.000159(33)
lowest	0.00075	0.00351(67)	0.00528(66)	0.000152(32)
order	0.0005	0.00156(69)	0.00270(68)	0.000065(33)
algorithm	0.00025	−0.00029(80)	0.00027(78)	−0.000018(38)
	0.0001	0.00094(80)	0.00109(79)	0.000043(38)
	0.00005	0.00056(58)	0.00068(57)	0.000026(27)
	0.001	0.00008(74)	0.00001(72)	−0.000014(35)
improved	0.00075	0.00079(60)	0.00049(59)	0.000022(28)
algorithm	0.0005	−0.00017(74)	−0.00036(71)	−0.000024(35)
	0.00025	0.0004(12)	0.0001(12)	0.000001(58)
	0.0001	0.00053(69)	0.00026(68)	0.000010(33)

Table 1: Stepsize dependence for $\mu = 1$ (disordered phase), $\beta = 0.125$ and $h = 0.02$ on a lattice of size 10^3.

	ϵ	$\langle \mathrm{Tr}\, U \rangle$	$\langle \mathrm{Tr}\, U^{-1} \rangle$	$\langle n \rangle$
	0.001	1.69646(39)	1.73658(35)	0.67976(16)
lowest	0.00075	1.69872(35)	1.73883(31)	0.68066(14)
order	0.0005	1.70165(40)	1.74158(35)	0.68184(16)
algorithm	0.00025	1.70475(46)	1.74450(41)	0.68308(18)
	0.0001	1.70442(36)	1.74427(32)	0.68295(15)
	0.00005	1.70586(42)	1.74561(37)	0.68352(17)
	extrapolation	1.70615(27)	1.74590(24)	0.68364(11)
	0.001	1.70605(40)	1.74571(36)	0.68360(16)
improved	0.00075	1.70514(27)	1.74486(24)	0.68324(11)
algorithm	0.0005	1.70629(43)	1.74598(38)	0.68370(17)
	0.00025	1.70597(29)	1.74571(26)	0.68357(12)
	0.0001	1.70596(39)	1.74576(35)	0.68356(16)
	average	1.70576(15)	1.74549(13)	0.683485(60)
	ϵ	$\langle L\mathrm{Tr}\, U \rangle$	$\langle L\mathrm{Tr}\, U^{-1} \rangle$	$\langle Ln \rangle$
	0.001	0.0480(10)	0.05258(97)	0.01922(40)
lowest	0.00075	0.03198(86)	0.03488(85)	0.01281(34)
order	0.0005	0.01944(98)	0.02119(96)	0.00779(39)
algorithm	0.00025	0.0084(14)	0.0092(13)	0.00335(55)
	0.0001	0.0055(11)	0.0058(10)	0.00219(42)
	0.00005	0.0020(12)	0.0021(12)	0.00081(50)
	0.001	0.0070(11)	0.0056(11)	0.00189(43)
improved	0.00075	0.00624(88)	0.00443(85)	0.00169(35)
algorithm	0.0005	0.0044(12)	0.0026(12)	0.00096(48)
	0.00025	0.00285(67)	0.00100(65)	0.00040(27)
	0.0001	0.0020(10)	0.0001(10)	0.00006(42)

Table 2: As in Table 1, for $\mu = 3$ (ordered phase).

Chapter 7

Conclusion

7.1 Problems with complex Langevin dynamics

The main problems with complex Langevin dynamics have long been known to be the issue of runaway solutions in numerical simulations and the more serious problem of convergence to the wrong limit. Of these, the first problem, of runaways, can be cured with an adaptive stepsize. The algorithm for the implementation of such a scheme has been outlined in this document and has been shown to be sufficiently robust to prevent runaways in all models examined thus far and, crucially, the simulations do not get stuck. The algorithm is also general to non-abelian groups, so that the same adaptive stepsize algorithm can be employed in, for instance, both $U(1)$ and $SU(3)$ models with no loss of generality. The adaptive stepsize algorithm can also, with little modification, be applied to the improved integration algorithm described in this work.

The adaptive stepsize algorithm can be recast as the introduction of a kernel. However, in order for the associated Fokker-Planck equation to have the same equilibrium solution, an extra term must be included into the Langevin

equation. In a field theory, this extra term is only relevant to a single site and its neighbours, therefore its effect is suppressed by the inverse of the lattice volume. However, it is less clear that neglecting the term in simple models of a single degree of freedom is correct. Nevertheless, empirical studies, such as the SU(3) toy model considered in Chapter 6, produce correct results.

Convergence to the wrong limit remains a more vexing problem. The XY model exhibits this behaviour in a peculiar manner: the complex Langevin results are correct in the region of large β, which corresponds to the part of the phase diagram where the nearest neighbour spins are correlated (the ordered phase). As the coupling β is reduced, the complex Langevin simulation converges to an incorrect distribution and gives wrong results. Importantly, it is the part of the phase diagram which dictates whether correct results are produced or not, rather than the severity of the sign problem. This implies that it is a property of the disordered phase, and therefore the physics, of the XY model which is causing the simulation to give incorrect results. This is distinct from other approaches at finite density.

An important test for convergence is to identify whether the simulation is attracted to the real axis when $\mu = 0$. Although in this there is no sign problem, by comparing two simulations in which the initial conditions are either on the real axis or spread in the complex plane, it is possible to identify whether the complex Langevin dynamics is restoring to the real axis. If the simulation which starts in the complex plane does not restore to the real axis then it is clear that the complex simulation is converging incorrectly and will give wrong results. With the XY model in the disordered phase, if even a single angle has initially an imaginary part, the simulation does not converge to the real axis at $\mu = 0$. Instead, the thermalised distribution is very wide in the imaginary direction with slowly decaying tails. The formal justification of complex Langevin dynamics show that this is indicative of non-vanishing boundary terms and therefore incorrect convergence. However, the SU(3) spin model converged back to the real axis at $\mu = 0$ in both phases and did

not display any slow decay in histograms.

7.2 Success with complex Langevin dynamics

The evidence from numerical studies shows that the failure of complex Langevin dynamics is independent of the strength of the sign problem, which means that the problem must have a different origin [77]. This work considers two models: the XY model, which suffers from incorrect convergence, and the SU(3) spin model, for which complex Langevin dynamics works without problem. From these studies, we have developed a set of tests which a correctly converging complex Langevin process must satisfy and have shown that the origin of incorrect convergence enters from distributions which are wide in the imaginary direction.

When the chemical potential is small, methods other than complex Langevin can be applied so that the issue of incorrect convergence can be tested against reliable results in this region. The two main methods which can be used here are imaginary chemical potential and Taylor series expansion.

In order for the complex Langevin process to converge correctly for all μ, it is necessary for the process to converge correctly in the special case of $\mu = 0$. Note that in this case the action is real, meaning that complex Langevin is strictly speaking not required. However, by starting with an initial configuration which is in the complexified space, the dynamics will nevertheless be complex. Since there is no chemical potential, the complex Langevin process must relax to the real axis and same distribution as the real Langevin process. Unless the hot start process converges correctly when $\mu = 0$, it will definitely not converge correctly when $\mu \neq 0$ and the action is complex. Indeed, in the case of the XY model, the hot and cold starts do not match, meaning that the complex Langevin process does not converge to the correct distribution.

Ideally, one would like to know whether a complex Langevin simulation

of a full lattice field theory will converge to the correct distribution and give trustworthy results. Since for theories of interest the correct results will not yet be known, one would ideally like a quantitative test to ascertain the validity of the results of the complex Langevin simulation. Simple toy models, in which the nearest neighbours of the full theory are written as arbitrary complex parameters, can be used for such a test. Since these models will generally be zero-dimensional, the correct results can be calculated directly. By scanning over a wide grid of complex parameters and comparing with the exact results, it is possible to see whether complex Langevin dynamics converge correctly. Indeed, correct convergence for all complex parameters guarantees that the full theory will also converge correctly. One can see this by considering each site coming into equilibrium with a heat bath that represents its neighbours. However, this is perhaps too strong a requirement, since it is possible that some complex parameters, which are actually combinations of the nearest neighbour interactions, will not occur in the full theory because of the physics of the theory.

One of the major problems with the complex Langevin method is that the formal proofs which show real Langevin dynamics converges to the correct distribution cannot be applied when the action is complex. However, as outlined in Chapter 5, the method can be justified by comparing the evolution of a complex density which represents the original problem and a real and positive density which represents the complex Langevin process. The argument shows that provided the decay of the real density is sufficiently fast to allow a partial integration without boundary terms, then the complex Langevin process is justified, meaning that it's results are correct.

The formal justification also provides a direct test for correct convergence. From the requirement that the boundary terms on the integration by parts vanish, it follows that a set of identities also vanish. These can be checked numerically in a complex Langevin simulation. This means that when the complex Langevin simulation goes wrong, the symptoms can be identified.

This means that when the simulation produces wrong results, the tests and criteria thus developed can be used to identify when this occurs.

Although the probability distribution cannot generally be computed directly from the Fokker-Planck equations, it is possible to compute a histogram of the configurations generated by the complex Langevin process. By analysis the decay of this, one can infer whether the criteria for the justification are satisfied.

In conclusion, the complex Langevin method offers great hope for simulating theories with a complex action and a severe sign problem. Although the problem of incorrect convergence as yet does not have a cure, the tests and criteria for correctness discussed in this document provide the basis for a self consistent mechanism for validating results from complex Langevin dynamics. Furthermore, the insights which have been discussed in this document show that there is still much to be learnt about complex Langevin dynamics.

Bibliography

[1] P. de Forcrand, PoS **LAT2009** (2009) 010 [arXiv:1005.0539 [hep-lat]].

[2] Z. Fodor and S. D. Katz, Phys. Lett. B **534** (2002) 87 [hep-lat/0104001]; JHEP **0203** (2002) 014 [hep-lat/0106002]; JHEP **0404** (2004) 050 [hep-lat/0402006].

[3] Z. Fodor, S. D. Katz and K. K. Szabo, Phys. Lett. B **568** (2003) 73 [hep-lat/0208078].

[4] C. R. Allton *et al.*, Phys. Rev. D **66** (2002) 074507 [hep-lat/0204010]; Phys. Rev. D **68** (2003) 014507 [hep-lat/0305007]; Phys. Rev. D **71** (2005) 054508 [hep-lat/0501030].

[5] R. V. Gavai and S. Gupta, Phys. Rev. D **68** (2003) 034506 [hep-lat/0303013].

[6] P. de Forcrand and O. Philipsen, Nucl. Phys. B **642** (2002) 290 [hep-lat/0205016].

[7] M. D'Elia and M. P. Lombardo, Phys. Rev. D **67** (2003) 014505 [hep-lat/0209146].

[8] M. D'Elia and F. Sanfilippo, Phys. Rev. D **80** (2009) 111501 [0909.0254 [hep-lat]].

[9] S. Kratochvila and P. de Forcrand, PoS **LAT2005** (2006) 167 [hep-lat/0509143].

[10] A. Alexandru, M. Faber, I. Horvath and K. F. Liu, Phys. Rev. D **72** (2005) 114513 [hep-lat/0507020].

[11] S. Ejiri, Phys. Rev. D **78** (2008) 074507 [0804.3227 [hep-lat]].

[12] Z. Fodor, S. D. Katz and C. Schmidt, JHEP **0703** (2007) 121 [hep-lat/0701022].

[13] K. N. Anagnostopoulos and J. Nishimura, Phys. Rev. D **66** (2002) 106008 [hep-th/0108041].

[14] J. Ambjorn, K. N. Anagnostopoulos, J. Nishimura and J. J. M. Verbaarschot, JHEP **0210** (2002) 062 [hep-lat/0208025].

[15] G. Akemann, J. C. Osborn, K. Splittorff and J. J. M. Verbaarschot, Nucl. Phys. B **712** (2005) 287 [hep-th/0411030].

[16] K. Splittorff, hep-lat/0505001.

[17] K. Splittorff and J. J. M. Verbaarschot, Phys. Rev. Lett. **98** (2007) 031601 [hep-lat/0609076].

[18] J. Han and M. A. Stephanov, Phys. Rev. D **78** (2008) 054507 [0805.1939 [hep-lat]].

[19] J. C. R. Bloch and T. Wettig, JHEP **0903**, 100 (2009) [0812.0324 [hep-lat]].

[20] M. P. Lombardo, K. Splittorff and J. J. M. Verbaarschot, Phys. Rev. D **80** (2009) 054509 [0904.2122 [hep-lat]]; Phys. Rev. D **81** (2010) 045012 [0910.5482 [hep-lat]].

[21] J. Danzer, C. Gattringer, L. Liptak and M. Marinkovic, Phys. Lett. B **682** (2009) 240 [0907.3084 [hep-lat]].

[22] J. O. Andersen, L. T. Kyllingstad and K. Splittorff, JHEP **1001**, 055 (2010) [0909.2771 [hep-lat]].

[23] S. Hands, T. J. Hollowood and J. C. Myers, 1003.5813 [hep-th].

[24] B. Bringoltz, 1004.0030 [hep-lat].

[25] J. Ambjorn, K. N. Anagnostopoulos, J. Nishimura and J. J. M. Verbaarschot, JHEP **0210** (2002) 062 [hep-lat/0208025].

[26] S. Ejiri, Phys. Rev. D **77** (2008) 014508 [arXiv:0706.3549 [hep-lat]].

[27] K. N. Anagnostopoulos, T. Azuma and J. Nishimura, Phys. Rev. D **83** (2011) 054504 [arXiv:1009.4504 [cond-mat.stat-mech]].

[28] S. Chandrasekharan and U. J. Wiese, Phys. Rev. Lett. **83** (1999) 3116 [cond-mat/9902128].

[29] M. G. Endres, Phys. Rev. D **75** (2007) 065012 [hep-lat/0610029].

[30] S. Chandrasekharan, PoS **LATTICE2008** (2008) 003 [0810.2419 [hep-lat]].

[31] D. Banerjee and S. Chandrasekharan, 1001.3648 [hep-lat].

[32] S. Chandrasekharan and U. -J. Wiese, Phys. Rev. Lett. **83** (1999) 3116 [arXiv:cond-mat/9902128 [cond-mat.stat-mech]].

[33] J. Bloch, Phys. Rev. Lett. **107** (2011) 132002 [arXiv:1103.3467 [hep-lat]].

[34] P. de Forcrand and M. Fromm, Phys. Rev. Lett. **104** (2010) 112005 [arXiv:0907.1915 [hep-lat]].

[35] J. Langelage, S. Lottini and O. Philipsen, JHEP **1102** (2011) 057 [Erratum-ibid. **1107** (2011) 014] [arXiv:1010.0951 [hep-lat]].

[36] Y. D. Mercado, H. G. Evertz and C. Gattringer, Phys. Rev. Lett. **106** (2011) 222001 [arXiv:1102.3096 [hep-lat]].

[37] C. Gattringer, Nucl. Phys. B **850** (2011) 242 [arXiv:1104.2503 [hep-lat]].

[38] Y. Delgado, H. G. Evertz, C. Gattringer and D. Goschl, arXiv:1111.0916 [hep-lat].

[39] W. Unger and P. de Forcrand, arXiv:1111.1434 [hep-lat].

[40] M. Fromm, J. Langelage, S. Lottini and O. Philipsen, arXiv:1111.4953 [hep-lat].

[41] O. Philipsen, arXiv:1111.5370 [hep-ph].

[42] G. Parisi, Phys. Lett. **B131** (1983) 393-395.

[43] J. R. Klauder, Stochastic quantization, in: H. Mitter, C.B. Lang (Eds.), Recent Developments in High-Energy Physics, Springer-Verlag, Wien, 1983, p. 351; J. Phys. A: Math. Gen. **16** (1983) L317; Phys. Rev. A **29** (1984) 2036.

[44] J. R. Klauder and W. P. Petersen, J. Stat. Phys. 39 (1985) 53.

[45] F. Karsch, H. W. Wyld, Phys. Rev. Lett. **55** (1985) 2242.

[46] J. Ambjorn and S. K. Yang, Phys. Lett. B **165** (1985) 140.

[47] J. Ambjorn, M. Flensburg and C. Peterson, Nucl. Phys. B **275** (1986) 375.

[48] J. Flower, S. W. Otto and S. Callahan, Phys. Rev. D **34**, 598 (1986).

[49] E. M. Ilgenfritz, Phys. Lett. B **181**, 327 (1986).

[50] P. H. Damgaard and H. Hüffel, Phys. Rept. **152** (1987) 227.

[51] J. Berges and I.-O. Stamatescu, Phys. Rev. Lett. **95** (2005) 202003 [hep-lat/0508030].

[52] J. Berges, S. Borsanyi, D. Sexty and I. O. Stamatescu, Phys. Rev. D **75** (2007) 045007 [hep-lat/0609058].

[53] J. Berges and D. Sexty, Nucl. Phys. B **799** (2008) 306 [0708.0779 [hep-lat]].

[54] G. Aarts and I.-O. Stamatescu, JHEP **0809** (2008) 018 [0807.1597 [hep-lat]].

[55] G. Aarts, Phys. Rev. Lett. **102** (2009) 131601 [0810.2089 [hep-lat]].

[56] G. Aarts, JHEP **0905** (2009) 052 [0902.4686 [hep-lat]].

[57] G. Aarts, F. A. James, E. Seiler, I. -O. Stamatescu, Phys. Lett. **B687** (2010) 154 [arXiv:0912.0617 [hep-lat]].

[58] G. Aarts, E. Seiler, I. -O. Stamatescu, Phys. Rev. **D81** (2010) 054508 [arXiv:0912.3360 [hep-lat]].

[59] C. Pehlevan and G. Guralnik, Nucl. Phys. B **811**, 519 (2009) [0710.3756 [hep-th]],

[60] G. Guralnik and C. Pehlevan, Nucl. Phys. B **822** (2009) 349 [0902.1503 [hep-lat]].

[61] H. Nakazato and Y. Yamanaka, Phys. Rev. D **34** (1986) 492.

[62] P. Hasenfratz and F. Karsch, Phys. Lett. B **125** (1983) 308.

[63] M. Campostrini, M. Hasenbusch, A. Pelissetto, P. Rossi and E. Vicari, Phys. Rev. B **63** (2001) 214503 [cond-mat/0010360].

[64] A. Roberge and N. Weiss, Nucl. Phys. B **275**, 734 (1986).

[65] P. de Forcrand and O. Philipsen, Nucl. Phys. B **642** (2002) 290 [hep-lat/0205016].

[66] R. De Pietri, A. Feo, E. Seiler and I. O. Stamatescu, Phys. Rev. D **76** (2007) 114501 [0705.3420 [hep-lat]].

[67] W. P. Petersen, hep-lat/9602008.

[68] N. Prokof'ev and B. Svistunov, Phys. Rev. Lett. **87** (2001) 160601.

[69] T. Banks, R. Myerson and J. B. Kogut, Nucl. Phys. B **129** (1977) 493.

[70] U. Wenger, Phys. Rev. D **80**, 071503 (2009) [0812.3565 [hep-lat]].

[71] P. de Forcrand and M. Fromm, Phys. Rev. Lett. **104** (2010) 112005 [0907.1915 [hep-lat]].

[72] U. Wolff, Nucl. Phys. B **824** (2010) 254 [Erratum-ibid. **834** (2010) 395] [0908.0284 [hep-lat]].

[73] J. Langelage and O. Philipsen, JHEP **1004** (2010) 055 [1002.1507 [hep-lat]].

[74] T. D. Cohen, Phys. Rev. Lett. **91**, 222001 (2003) [hep-ph/0307089].

[75] G. Aarts, unpublished (2009).

[76] G. Aarts, E. Seiler and I. O. Stamatescu, in progress.

[77] G. Aarts, F. A. James, E. Seiler, I. -O. Stamatescu, Eur. Phys. J. **C71** (2011) 1756 [arXiv:1101.3270 [hep-lat]].

[78] F. Karsch and H. W. Wyld, Phys. Rev. Lett. **55** (1985) 2242.

[79] N. Bilic, H. Gausterer and S. Sanielevici, Phys. Rev. D **37** (1988) 3684.

[80] G. Aarts, E. Seiler and I. -O. Stamatescu, Phys. Rev. D **81** (2010) 054508 [arXiv:0912.3360 [hep-lat]].

[81] G. Aarts and F. A. James, JHEP **1008** (2010) 020 [arXiv:1005.3468 [hep-lat]].

[82] G. Aarts and K. Splittorff, JHEP **1008** (2010) 017 [arXiv:1006.0332 [hep-lat]].

[83] G. Aarts, F. A. James, E. Seiler and I. -O. Stamatescu, Eur. Phys. J. C **71** (2011) 1756 [arXiv:1101.3270 [hep-lat]].

[84] Chien-Cheng Chang, Math. Comp. **49** 180 (1987) 523-542.

[85] J. C. Osborn, K. Splittorff and J. J. M. Verbaarschot, Phys. Rev. Lett. **94**, 202001 (2005) [hep-th/0501210].

[86] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer Verlag (1992).

[87] I. T. Drummond, S. Duane and R. R. Horgan, Nucl. Phys. B **220** (1983) 119.

[88] S. M. Catterall, I. T. Drummond and R. R. Horgan, Phys. Lett. B **254** (1991) 177.

[89] K. Okano, L. Schulke and B. Zheng, Prog. Theor. Phys. Suppl. **111** (1993) 313.

[90] M. -P. Lombardo, Nucl. Phys. Proc. Suppl. **83** (2000) 375 [hep-lat/9908006].

[91] D. Banerjee and S. Chandrasekharan, Phys. Rev. D **81** (2010) 125007 [arXiv:1001.3648 [hep-lat]].