



THE UNIVERSITY *of* EDINBURGH
School of Physics
and Astronomy

Fourier Acceleration of the $SU(2) \otimes SU(2)$ Principal Chiral Model in 2 Dimensions

MPhys Project Report

Julian Wack

Submitted for the 40pt MPhys Project course PHYS11016
March 26, 2023

Abstract

Motivated by the similarity to QCD, specifically the property of asymptotic freedom, we simulate the dynamics of the simplest principal chiral model in 2 dimensions using the Hamiltonian Monte Carlo algorithm. By introducing Fourier Acceleration, we increase the simulation efficiency by over a factor of 300. This yields numerical predictions at a precision exceeding that of existing studies and allows us to verify the onset of asymptotic scaling. The analysis and simulation implementation form a foundation for future studies of more complex principal chiral models.

Supervisors: Dr Brian Pendleton, Dr Roger Horsley

Personal statement

The largest part of my project work was related to writing my own simulation routines in Python and subsequently testing and debugging these. Throughout the project, I wrote a primary version of this report which helped my own understanding while also allowing me to track progress and failures.

To familiarise myself with the field of lattice simulations and the Hamiltonian Monte Carlo algorithm, I studied previous student projects on simulating the one-dimensional quantum harmonic oscillator (QHO). Within the first 3 weeks, I produced my own implementation and subsequently turned my attention to adding Fourier Acceleration, which has not been done previously.

In parallel, I developed a hyperparameter calibration procedure which allows for better automation and more consistent comparisons between simulations. Over the course of roughly 2 weeks, I researched and implemented an efficient method to compute correlations based on Fast Fourier transforms. The final routine is widely employable and heavily used for the computation of correlation functions as well as the measurement error estimation.

Towards the middle of November, I had completed the study of the QHO and moved towards the main theory of this report, the $SU(2) \otimes SU(2)$ model in two dimensions. Due to the complexity of the model, and Python code being slow, I spent the initial part of this project-phase optimising matrix operations for $SU(2)$ elements. Reusing parts of the QHO implementation allowed me to run the first $SU(2) \otimes SU(2)$ simulations near the end of semester 1.

Over the Christmas break, I initially focused on studying strong and weak coupling expansions and was able to reproduce numerical results in the literature. Afterwards, my attention shifted towards computing the wall-to-wall correlation function. Here I noticed an opportunity to employ the efficient algorithm developed during the QHO study. In parallel, I worked on the Fourier Acceleration of this model and arrived at a working implementation in the second week of semester 2.

Until the end of February, I addressed the two main objectives of the project: the numerical investigation of asymptotic freedom and measuring to what degree critical slowing down can be avoided through employing Fourier Acceleration. The interplay of statistical and systematic errors when determining the correlation length as well as the necessity of large statistics made achieving the former goal difficult and time intensive. I was able to make a breakthrough towards the second objective by resolving a bottleneck in the computation of the integrated autocorrelation time, speeding up the analysis by several orders of magnitude.

For the production runs I moved my computations to the student computing cluster. In the process, it became apparent that due to a design flaw in my simulations, an unfeasible amount of memory was required for long runs. Changing the code structure seemed the only available solution and delayed the main data-taking phase by close to 2 weeks. After extensive tests, I submitted the production runs, the longest of which ran for over 2 weeks.

Checking for finite size effects and testing the robustness of the analysis routines required the submission of further jobs. From mid-March onwards, I focused on interpreting the collected data and writing up my results.

Acknowledgments

Thank you to both my supervisors, Dr Brian Pendleton and Dr Roger Horsley, for introducing me to the exciting domain of lattice field theory. Their guidance, support, and encouragement throughout the project allowed me to achieve the best results possible, and for that I am grateful.

Contents

1	Introduction	1
2	Background	2
2.1	Lattice Field Theory	2
2.2	Markov Chain Monte Carlo	3
2.3	Hamiltonian Monte Carlo	5
2.4	Numerical Implementation	8
3	Error Estimation and Autocorrelation	10
3.1	Autocorrelation Function and Time	10
3.2	Efficient Computation	11
4	Fourier Acceleration	12
4.1	Critical Slowing Down	12
4.2	Modified Dynamics	13
4.3	Acceleration in Practice	14
4.4	Distribution of Momenta	16
4.5	The Fourier Accelerated Hamiltonian Monte Carlo Algorithm	18
5	Quantum Harmonic Oscillator	19
5.1	Measurement Results	19
5.2	Fourier Acceleration	22

6	$SU(2) \otimes SU(2)$ Chiral Model in 2 Dimensions	24
6.1	Motivation	24
6.2	$SU(2)$ Matrix Properties	25
6.3	Action	26
6.4	Equations of Motion	27
6.4.1	Field	27
6.4.2	Momenta	28
6.4.3	Discretisation	29
6.5	Fourier Acceleration	29
6.5.1	Simple Simulation Check	30
6.6	Strong and Weak Coupling Expansions	31
6.7	Correlations	33
6.7.1	Wall to Wall correlations	33
6.7.2	Effective Mass	35
6.8	The β Function and Asymptotic Scaling	36
6.8.1	Mass over Λ Ratio	37
6.9	Critical Slowing Down	39
7	Conclusion	43
	References	44
	Appendices	47
A	Additional Proofs	47
A.1	Liouville's Theorem	47
A.2	Cross Correlation Theorem for Discrete Variables	47
B	Quantum Harmonic Oscillator	48
B.1	Analytic Results	48
B.2	Numerical Ground State Results	48
C	$SU(2) \otimes SU(2)$	50
C.1	Coupling Expansion Residuals	50
C.2	Improved Coupling	50

1 Introduction

The Standard Model of particle physics is an extremely powerful theory which, to this date, yields predictions consistent with experimental observation at the 5σ level. However, a wealth of evidence for physics beyond the Standard Model exists and many fundamental questions are left unanswered by the theory. A non-exhaustive list includes confinement, the values of the CKM matrix as well as dark matter and the hierarchy problem referring to the vastly different strengths of the fundamental forces. In addition to direct searches for new physics at high energies, for example at collider experiments, one can also look at its indirect effects at low energy. In this regime, QCD becomes strongly coupled, causing perturbation theory to lose applicability. Some progress can be made when resorting to perturbation theory of effective theories for QCD. These, however, suffer from being non-renormalisable in general which causes difficulty in extracting results beyond leading order accuracy, necessary for meaningful comparison with experimental measurements [1].

Lattice field theory is a numerical approach allowing to treat the system in question fully non-perturbatively and is capable of producing high-precision predictions in regimes largely inaccessible to theoretical computations for the aforementioned reasons. A prominent example where the three main inference methods of particle physics (experimental measurement, theoretical calculation, and lattice computation) work together to potentially reveal new physics is the anomalous magnetic moment $g-2$ of the muon. The theoretical result relies on the input of the R-ratio data for $e^+e^- \rightarrow \text{hadrons}$ which, when taken from lattice QCD results rather than experimental data, results in a more significant disagreement with the measured value of $g-2$ and is currently placed at the 4.2σ level [2, 3, 4].

The above example illustrates the importance of high-precision lattice computations. These require the lattice used to approximate continuous space-time to have a fine spacing and to be as large as possible. However, taking these processes to the extreme is known to cause a rapid simulation efficiency decrease. It is this diverging simulation cost, commonly referred to as critical slowing down, that motivates the algorithmic segment of this report.

In particular, we consider the simplest member of the principal chiral quantum field theories $SU(N) \otimes SU(N)$ as these are known to share the property of asymptotic freedom with non-abelian gauge theories such as QCD while being computationally easier to work with [5]. For the $N = 2$ model, we seek to measure the ability of a technique referred to as Fourier Acceleration to avoid critical slowing down. In addition to this algorithmic objective, we also aim to close a gap in the literature by numerically verifying asymptotic freedom. Such studies exist for some selected values of N [5, 6, 7], but not the $N = 2$ case. We hope that the techniques developed in this report provide a fruitful foundation for the analysis of more complicated models and eventually towards understanding QCD better.

The remaining project is structured as follows. In sec. 2, we present an introduction to Markov chain Monte Carlo methods which leads on to discussing the Hamiltonian Monte Carlo algorithm. In order to address the algorithmic goal of the project, autocorrelation times and their efficient computation are discussed in sec. 3. This is followed by a general discussion on Fourier Acceleration and how it can resolve the issue of critical slowing down.

Subsequently, the introduced techniques are applied to two systems. In sec. 5, selected results of the quantum harmonic oscillator study are presented. The main results of this report are contained in sec. 6. We show the onset of asymptotic scaling but conclude that improved couplings techniques are required to make more robust claims. We tested one such prescription without obtaining any improvements. Finally, we find that the simulation efficiency close to the continuum can be improved by a factor 320 when employing Fourier Acceleration and conclude the study in sec. 7.

The simulation code and the collected data are publicly available at [8].

2 Background

2.1 Lattice Field Theory

Suppose we seek to approximate the expectation value for some observable O in a theory involving some generic field ϕ and characterised by the action $S[\phi]$ in Minkowski space. In field theory, the expectation is defined through the path integral according to

$$\langle O(\phi) \rangle = \int \mathcal{D}\phi O(\phi) \exp(iS[\phi]). \quad (1)$$

An equivalent formulation is obtained by performing a Wick rotation to imaginary time. Via $t \rightarrow it$, this transforms the Minkowski metric into a Euclidean one

$$x^2 = x^\mu x^\nu \eta_{\mu\nu} = t^2 - \mathbf{x}^2 \rightarrow -t^2 - \mathbf{x}^2 = -x^\mu x^\nu \delta_{\mu\nu} \quad (2)$$

and further yields

$$S[\phi] = \int d^4x \mathcal{L}[\phi] \rightarrow i \int d^4x \mathcal{L}_E[\phi] = iS_E[\phi]. \quad (3)$$

The subscript E indicates that the quantity is now defined in the Euclidean sense but will be dropped from here on as we will exclusively work in the Euclidean formulation of the path integral. The advantage of this analytic continuation in the time variable is that the expectation now takes the more familiar form of a statistical system. Specifically

$$\langle O(\phi) \rangle = \int \mathcal{D}\phi O(\phi) \exp(-S[\phi]) \quad (4)$$

which allows to identify $P(\phi) \equiv \exp(-S[\phi])$ as the probability density of the field ϕ . Most systems of interest are those for which the integration cannot be performed analytically. Instead, we seek to approximate it through Monte Carlo techniques. In order to do so, one must abandon continuous space-time and resort to defining the system on a space-time lattice with finite side length L and finite spacing a . Otherwise, the integral in eq. 4 is infinite-dimensional and therefore cannot be approximated numerically. An important feature of lattice discretization is the natural regularisation of infrared and ultraviolet divergences through the finite lattice size (long distance cutoff) and finite lattice spacing (short distance cutoff) respectively [1, 9].

However, through this process, two types of errors are introduced. Firstly, a statistical error originating from the Monte Carlo approximation which can be systematically reduced by running longer simulations and is discussed in more detail in sec. 3. The second error is a systematic one and due to approximating a continuous, infinite volume by a discrete, finite lattice. In order to obtain predictions comparable to experimental measurements, simulations should be performed using a large and finely spaced lattice. Again, being able to adjust these simulation parameters makes this error controllable but results in a diverging simulation cost as critical slowing down occurs.

2.2 Markov Chain Monte Carlo

The essence behind numerically estimating the expectation in eq. 4 lies in constructing a stochastic estimator based on samples $\{\phi_i\}$ which are drawn independently from the probability distribution $P(\phi)$ [10]. The point that these must be independent is crucial and the key underlying issue of critical slowing down. For a total of M such samples, the Monte Carlo estimator is given by

$$\langle O \rangle \approx \bar{O} \equiv \frac{1}{M} \sum_{i=1}^M O_i \quad (5)$$

which asymptotically converges to $\langle O \rangle$ as $M \rightarrow \infty$. The great advantage of Monte Carlo techniques is that the estimator error is $\mathcal{O}(1/\sqrt{M})$ and thus independent of the dimensionality of the target distribution $P(\phi)$ [11]. This is of key importance since we seek to perform simulations close to the continuum limit and each further lattice site contributes additional degrees of freedom which in turn increase the dimensionality of $P(\phi)$.

In the context of lattice simulations, a sample refers to a particular field configuration defined on the lattice. These configurations will be collected using a Markov chain, referring to a sequence of random variables where the next element ϕ^* conditionally depends on the current one ϕ . Such a step in the chain is taken according to the transition probability $T(\phi^*|\phi)$, which, when chosen suitably, causes the chain to converge on a state space region, known as the typical set. In this region the probability mass $P(\phi)\mathcal{D}\phi$ is sharply peaked and thus provides the dominant contributions to the expectation in eq. 4. As the dimensionality of the target distribution increases, the typical set becomes increasingly concentrated causing contributions from other regions to become negligible [12]. In more detail, the development of the chain can be broken down into 3 stages, schematically displayed in fig. 1.

1. Convergence: After initialising the chain with an arbitrary starting point ϕ_0 , the next few configurations explore the state space and gradually approach the typical set.
2. Exploration: Configurations are now drawn dominantly from the typical set, causing the precision of the running average of eq. 5 to improve quickly.
3. Equilibrium: The configurations have fully explored the typical set, such that all details of the target distribution are encapsulated in $\{\phi\}$.

It is evident that a number of steps in the chain are necessary before the collected configurations accurately represent $P(\phi)$, at which point the chain is said to have thermalised or equilibrated. All configurations prior to this should not be used in the estimation of the expectation value and are therefore rejected as burn in. Declaring the first 10% of the collected samples as burn in has become conventional in the field but it is generally advised to inspect the chain of measurements manually as the rate of convergence differs between observables. Equilibrium is said to be reached once the measurements have approached a constant within statistical fluctuations.

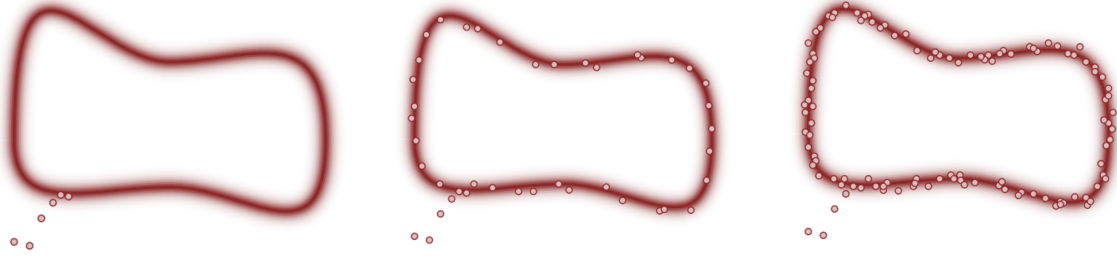


Figure 1. Illustration of the state space and typical set of some arbitrary target distribution, with the colour intensity describing the concentration of the probability mass. The different stages of populating the Markov chain are displayed from left to right with samples represented by dots. Initially, the chain moves from the arbitrary starting point towards the typical set and subsequently explores it. Finally, the samples populate the entire set which asymptotically reduces the approximation error in eq. 5 to zero. Taken from [12].

In order to achieve this desired behaviour, the transition probability must satisfy two conditions [13]:

- **Invariance of the target distribution:** If the current configuration ϕ is a state of $P(\phi)$, so must be the next configuration ϕ^* . This assures that the chain explores the typical set and does not diverge from it. More precisely, we require

$$P(\phi^*) = \int \mathcal{D}\phi T(\phi^*|\phi)P(\phi). \quad (6)$$

A condition more familiar to physicists (which is sufficient but not necessary) is the detailed balance constraint

$$T(\phi^*|\phi)P(\phi) = T(\phi|\phi^*)P(\phi^*). \quad (7)$$

Through the unitary requirement of the transition probability, eq. 6 is quickly verified.

- **Ergodicity:** Any state ϕ permitted by the probability density, meaning $P(\phi) > 0$, must be reachable by the chain in a finite number of steps. This assures that the chain converges to the target distribution regardless of the starting point.

Finally, in order to actually construct a Markov Chain, the issue of how to choose the transition probability must be addressed. Historically, the first such algorithm is the Metropolis

algorithm [14] which lays the groundwork for many more sophisticated prescriptions, including Hamiltonian Monte Carlo which will be discussed below. The algorithm can be separated into two steps: proposal and acceptance. First, generate a candidate configuration ϕ^* based on the current configuration ϕ according to some proposal distribution $q(\phi^*|\phi)$ ¹. The proposed candidate is accepted as the next configuration in the chain with probability

$$\alpha(\phi^*, \phi) = \min \left(1, \frac{q(\phi|\phi^*)P(\phi^*)}{q(\phi^*|\phi)P(\phi)} \right). \quad (8)$$

An acceptance probability of this form has the advantage that the normalisation of the target distribution cancels, which can be computationally expensive to evaluate. Should the candidate be rejected, ϕ is taken as the next configuration and thus repeated in the chain, corresponding to having a larger probability weight in the target distribution. The transition probability is consequently

$$T(\phi^*|\phi) = q(\phi^*|\phi)\alpha(\phi^*, \phi) \quad (9)$$

which can be easily shown to satisfy detailed balance:

$$T(\phi^*|\phi)P(\phi) = q(\phi^*|\phi)\alpha(\phi^*, \phi)P(\phi) \quad (10)$$

$$= \min(q(\phi^*|\phi)P(\phi), q(\phi|\phi^*)P(\phi^*)) \quad (11)$$

$$= q(\phi|\phi^*)\alpha(\phi, \phi^*)P(\phi^*) \quad (12)$$

$$= T(\phi|\phi^*)P(\phi^*). \quad (13)$$

The arguably simplest choice for $q(\phi^*|\phi)$ is a normal distribution, centred on ϕ with fixed covariance and yields the Random Walk Metropolis algorithm. Any distribution which is symmetric under $\phi \leftrightarrow \phi^*$ causes the acceptance probability to take an even simpler form, namely

$$\alpha(\phi^*, \phi) = \min \left(1, \frac{P(\phi^*)}{P(\phi)} \right). \quad (14)$$

This reveals some intuition behind the algorithm as configurations are favoured that yield a larger probability density and are thus located in the neighbourhood of the typical set. As with increasing dimensionality, the typical set becomes more and more concentrated, the random trial and error approach of the Random Walk Metropolis renders inefficient [12]. This motivates investigating alternative ways of proposing configurations which use the geometrical information of the typical set to make large steps in state space into unexplored regions of the typical set. A method which has proved to be incredibly successful at doing so is Hamiltonian Monte Carlo.

2.3 Hamiltonian Monte Carlo

Hamiltonian Monte Carlo (also known as Hybrid Monte Carlo and henceforth abbreviated by HMC) was originally introduced in [15] and has become an established method in the

¹In the original Metropolis algorithm, this distribution is assumed to be symmetric under exchanging $\phi \leftrightarrow \phi^*$. We will relax this constraint and therefore technically consider the Metropolis-Hastings algorithm.

lattice field theory community and beyond due to its simple extension of the Metropolis algorithm while providing significant efficiency improvements.

In the Metropolis and related algorithms, the typical set gets explored diffusively due to the stochastic nature of the proposal distribution. This causes distant, unsurveyed regions of the typical set to be reached slowly, yielding a non-optimal exploration overall. The solution proposed by HMC is to obtain candidates deterministically by introducing a fictitious time variable and simulating Hamiltonian dynamics through this time, using the current configuration as the initial condition [15]. Differential geometry allows to show that these trajectories have the property of tracing the typical set, such that the proposed candidate will have a high acceptance rate while being distant from the current configuration. The combination of these two properties results in the configurations of the chain being largely statistically independent and allows for a complete and efficient exploration of the typical set [12].

In order to exploit this property, momenta π conjugate to the field degrees of freedom must be introduced. This causes the state space (ϕ) to be extended to phase space (ϕ, π) , allowing us to construct a Hamiltonian and consequently deduce Hamilton's equations. Specifically,

$$H(\phi, \pi) = T(\pi) + S(\phi) \quad (15)$$

where T can be thought of as a kinetic energy depending on the momenta while the system's action takes the role of a potential term. The resulting Hamilton's equations are

$$\frac{d\phi}{dt} = \frac{\partial H}{\partial \pi}, \quad \frac{d\pi}{dt} = -\frac{\partial H}{\partial \phi}. \quad (16)$$

Starting from position (ϕ, π) in phase space and solving these equations over some period in the fictitious time, yields (ϕ^*, π^*) and thus acts as the proposal distribution $q(\phi^*|\phi)$ introduced in the previous section. Having laid out the logic behind HMC, it remains to show that this particular proposal choice satisfies the required properties to construct Markov chain updates. Proving the ergodicity of HMC is technically involved with proofs available in [10, 16].

Reversibility Hamiltonian dynamics is reversible, such that applying $t \rightarrow -t$, allows to recover (ϕ, π) from (ϕ^*, π^*) . Negating the endpoint momenta has the equivalent effect but is numerically easier to implement and thus commonly used in practice. It follows that

$$q(\phi^*, \pi^*|\phi, \pi) = q(\phi, -\pi|\phi^*, -\pi^*). \quad (17)$$

This symmetry causes the accept/reject update to reduce to eq. 14. In that sense, HMC can be viewed as an extension of the Metropolis algorithm with a particularly effective approach to proposing new candidates.

Liouville's theorem Let the evolution according to Hamilton's equations be denoted by the mapping $\mathcal{H} : (\phi, \pi) \rightarrow (\phi^*, \pi^*)$. It is a well-known property of Hamiltonian dynamics

that phase space volumes are preserved. This is a direct consequence of \mathcal{H} constituting a canonical transformation and thus obeying the symplectic condition [17]. For notational convenience, define $\eta = (\phi, \pi)$ and $\xi = (\phi^*, \pi^*)$ where ϕ and π are themselves a collection of f quantities. The Jacobian matrix associated with the mapping \mathcal{H} is then given by

$$M_{ij} \equiv \frac{\partial \xi_i}{\partial \eta_j} \quad \text{and satisfies} \quad M J M^T = J, \quad J = \begin{bmatrix} 0 & \mathbb{1}_{f \times f} \\ -\mathbb{1}_{f \times f} & 0 \end{bmatrix}. \quad (18)$$

Thus $\det M = \pm 1$ and, as Hamilton's equations are connected to the identity, we may discard the negative solution. Consequently, phase space volumes are preserved:

$$\mathcal{D}V^* = \mathcal{D}\phi^* \mathcal{D}\pi^* = \mathcal{D}\phi \mathcal{D}\pi \det M = \mathcal{D}\phi \mathcal{D}\pi = \mathcal{D}V. \quad (19)$$

In appendix A, we also present an alternative derivation of this theorem, not building symplecticness.

Building on these properties and following [15], we can show that the target distribution is left invariant by verifying the detailed balance in phase space:

$$\mathcal{D}\phi \mathcal{D}\pi T(\phi^*, \pi^* | \phi, \pi) P(\phi, \pi) = \mathcal{D}\phi^* \mathcal{D}(-\pi^*) T(\phi, -\pi | \phi^*, -\pi^*) P(\phi^*, -\pi^*). \quad (20)$$

Here, $P(\phi, \pi) = e^{-H(\phi, \pi)} = e^{-T(\pi)} e^{-S(\phi)}$ and choosing the kinetic energy quadratic in π results in the Hamiltonian, and thus α from eq. 14, to be invariant under the negation of the momenta. Let $H = H(\phi, \pi)$, $H^* = H(\phi^*, \pi^*)$ and $\Delta H = H^* - H$, such that we have

$$e^{-H} \alpha(\phi^*, \pi^*; \phi, \pi) = e^{-H} \min(1, e^{-\Delta H}) = e^{-H^*} \min(e^{\Delta H}, 1) = e^{-H^*} \alpha(\phi, \pi; \phi^*, \pi^*). \quad (21)$$

Using the definition of the transition rate and the above observation, followed by eq. 17 yields

$$T(\phi^*, \pi^* | \phi, \pi) P(\phi, \pi) = q(\phi^*, \pi^* | \phi, \pi) \alpha(\phi^* \pi^*; \phi, \pi) e^{-H} \quad (22)$$

$$= q(\phi^*, \pi^* | \phi, \pi) \alpha(\phi, \pi; \phi^*, \pi^*) e^{-H^*} \quad (23)$$

$$= q(\phi, -\pi | \phi^*, -\pi^*) \alpha(\phi, \pi; \phi^*, \pi^*) e^{-H^*}. \quad (24)$$

As H^* and α are independent of the sign of the momenta, we arrive at

$$T(\phi^*, \pi^* | \phi, \pi) P(\phi, \pi) = q(\phi, -\pi | \phi^*, -\pi^*) \alpha(\phi, -\pi; \phi^*, -\pi^*) e^{-H^*} \quad (25)$$

$$= T(\phi, -\pi | \phi^*, -\pi^*) P(\phi^*, -\pi^*). \quad (26)$$

Finally, due to Liouville's theorem and the invariance of the phase space volume elements under the negation of the momenta, the detailed balance condition is satisfied.

Observe that choosing the Hamiltonian of the form as in eq. 15 causes the phase space distribution to factorise into a ϕ and π dependent piece. Therefore, the field configurations follow the target distribution $e^{-S(\phi)}$ irrespective of the momenta such that they can be simply discarded after having fulfilled their purpose of facilitating Hamiltonian dynamics. Even though this underlines the auxiliary nature of the momenta, the choice of the kinetic energy does influence the efficiency of the algorithm. In this study, we employ the arguably

simplest choice of $T = \pi^2/2$, corresponding to the momenta distribution being a zero mean and unit variance Gaussian. This choice is part of a larger class known as Euclidean-Gaussian kinetic energies, allowing for a tune-able constant covariance which takes the role of the Euclidean metric. Generalising further by allowing the covariance to depend on ϕ , results in Riemannian Hamiltonian Monte Carlo which is one among many active areas of research [12] but will not be further investigated in this report. The full algorithm steps are illustrated in Algorithm 1.

Algorithm 1 Hamiltonian Monte Carlo

Input: Kinetic term $T(\pi)$ and action $S(\phi)$ such that $H(\phi, \pi) = T(\pi) + S(\phi)$.

Length τ of fictitious time to integrate Hamilton's equations over.

Length of Markov chain M .

Output: Markov chain $\{\phi_i\} = \{\phi_1, \dots, \phi_M\}$ with configurations following $P(\phi) = e^{-S(\phi)}$ once the chain has equilibrated.

- 1: Provide a starting configuration ϕ_0
 - 2: **for** $i = 1$ to M **do**
 - 3: Draw a momentum configuration: $\pi \sim e^{-T(\pi)}$
 - 4: Evolve Hamilton's equations for time τ : $(\phi, \pi) \rightarrow (\phi^*, \pi^*)$
 - 5: Reverse the candidate momentum: $\pi^* \rightarrow -\pi^*$
 - 6: Compute the acceptance probability: $\alpha = \min(1, e^{-(H(\phi^*, \pi^*) - H(\phi, \pi))})$
 - 7: Draw a uniform random number: $u \sim \text{Uniform}(0, 1)$
 - 8: **if** $\alpha > u$ **then**
 - 9: Accept the candidate as the new configuration: $\phi_{i+1} = \phi^*$
 - 10: **else**
 - 11: Reject the candidate and repeat the current configuration: $\phi_{i+1} = \phi$
 - 12: **end if**
 - 13: **end for**
-

The Metropolis update $\alpha = \min(1, e^{-(H(\phi^*, \pi^*) - H(\phi, \pi))})$ can be understood as accepting with certainty any candidates which lower the Hamiltonian, while the acceptance probability decays exponentially for those that cause $\Delta H > 0$. Since the considered Hamiltonian does not contain any explicit time dependence it is conserved along trajectories, resulting in $\alpha = 1$. In practice, Hamilton's equations are integrated numerically by discretizing the fictitious time variable t . This causes the analytically exact conservation to be broken, yielding $\alpha < 1$.

2.4 Numerical Implementation

The constraints of phase space preservation and reversibility, necessary to construct the Markov chain, limit the choice of a numerical integration scheme, with the class of symplectic integrators being commonly used. For a Euclidean-Gaussian kinetic energy, a particularly simple choice is the leapfrog scheme, which we will adopt.

Suppose we seek to evolve $(\phi, \pi) \rightarrow (\phi^*, \pi^*)$ by solving eq. 16 for time $\ell\epsilon$ using ℓ integration steps, each with a time step size ϵ . The leapfrog scheme attains its name from performing

alternating field and momenta updates with a time shift of $\epsilon/2$. The integration logic is the following:

1. Perform a half step for the momentum and a full step for the field:

$$\pi(\epsilon/2) = \pi(0) - \frac{\epsilon}{2} \frac{\partial S}{\partial \phi} \Big|_{t=0}, \quad \phi(\epsilon) = \phi(0) + \epsilon \frac{\partial T}{\partial \pi} \Big|_{t=\epsilon/2} \quad (27)$$

2. Perform $\ell - 1$ alternating full steps in the momentum and the field.

Let $n = 1, \dots, \ell - 1$:

$$\pi(n\epsilon + \epsilon/2) = \pi(n\epsilon - \epsilon/2) - \epsilon \frac{\partial S}{\partial \phi} \Big|_{t=n\epsilon}, \quad \phi(n\epsilon + \epsilon) = \phi(n\epsilon) + \epsilon \frac{\partial T}{\partial \pi} \Big|_{t=n\epsilon + \epsilon/2} \quad (28)$$

3. Perform a half step for the momentum:

$$\pi(\ell\epsilon) = \pi(\ell\epsilon - \epsilon/2) - \frac{\epsilon}{2} \frac{\partial S}{\partial \phi} \Big|_{t=\ell\epsilon} \quad (29)$$

and finally set $(\phi^*, \pi^*) = (\phi(\ell\epsilon), \pi(\ell\epsilon))$.

Apart from its simplicity, a crucial property of the leapfrog scheme is that it preserves phase space volumes for any step size ϵ . A proof is offered in [18]. That the reversibility requirement of eq. 17 is satisfied can be quickly verified by negating the momenta, applying the above evolution steps again and using the results from the forward evolution in reverse order. Specifically, the first step of the reverse evolution yields

$$\pi(\ell\epsilon + \epsilon/2) = -\pi(\ell\epsilon) - \frac{\epsilon}{2} \frac{\partial S}{\partial \phi} \Big|_{t=\ell\epsilon} = -\pi(\ell\epsilon - \epsilon/2) \quad (30)$$

$$\phi(\ell\epsilon + \epsilon) = \phi(\ell\epsilon) - \epsilon \frac{\partial T}{\partial \pi} \Big|_{t=\ell\epsilon + \epsilon/2} = \phi(\ell\epsilon). \quad (31)$$

Confirming the reversibility numerically offers a simple test that the leapfrog integration, as well as the involved derivatives, are implemented correctly.

During the leapfrog scheme, the hyperparameters ℓ, ϵ are introduced which increase the complexity of the model and require fine-tuning. A detailed discussion of the associated problems can be found in [10] while here we only give a high-level overview. The choice of ϵ determines the error of the numerical integration², resulting in $\Delta H > 0$ for any finite ϵ . Recalling that the acceptance rate α reduces with ΔH reveals that decreasing ϵ yields a larger acceptance rate at the price that the numerical integration becomes computationally more expensive for a fixed number of integration steps. The theoretical optimal value of the acceptance rate, resulting in a chain representing samples from the target distribution most accurately, can be shown to be $\alpha = 65\%$ [10]. In our simulations, we fix the trajectory

²It is shown in [19] that the global leapfrog error is $\mathcal{O}(\epsilon^2)$.

length to unity and calibrate ℓ, ϵ under this constraint until the acceptance rate lies close to its optimal value ($\alpha \in [0.6, 0.75]$). Adjusting the trajectory length offers a further handle on the algorithm's efficiency but will not be further explored in this project.

In this context, efficiency should be understood as a measure of the statistical independence between configurations in the chain. Independence is required in order for the chain to accurately represent samples of the target probability distribution. The next section discusses this in more detail through the notion of autocorrelations. In sec. 4 a modification to HMC known as Fourier Acceleration will be introduced which aims to resolve the issue of autocorrelations while keeping the trajectory length fixed.

3 Error Estimation and Autocorrelation

As described by eq. 5, once a the chain of configurations is constructed and a measurement of the observable O is taken for each sample, yielding the chain of observations $\{O_1, \dots, O_M\} = \{O_i\}$, the stochastic estimate for the expectation $\langle O \rangle$ is a simple average over this set. Having used a Markov chain to simulate the configurations makes estimating the error on \bar{O} less trivial as elements in $\{O_i\}$ are correlated. The standard error on the mean (SEM) is given by the standard deviation σ_O over the square root of the number of statistically independent observations $\sqrt{M'}$. Clearly applying this naively with $M' = M$ to $\{O_i\}$ will underestimate the true error of \bar{O} since the correlations in the chain cause $M = \tau M'$ where the integrated autocorrelation time (IAT) τ gives the average number of steps separating two statistically independent observations in $\{O_i\}$. Consequently, an improved error estimate, which we will refer to as the IAT corrected SEM and employ unless stated otherwise, is given by

$$\sigma_{\bar{O}} = \sqrt{\frac{\tau_O}{M}} \sigma_O, \quad (32)$$

with σ_O as the standard deviation of $\{O_i\}$ as before [20]. Note that in general, the IAT is specific to each observable, as indicated by the additional subscript in eq. 32. For notational convenience, this is left implied from here on.

Not only is the IAT necessary for reliable error estimation of observables, it also plays a central role in the algorithmic objective of this study, dealing with the scaling of the IAT as the lattice is taken to the continuum limit. It is therefore vital to be able to compute the IAT accurately and efficiently. This will be addressed in the remainder of this section and will be based on the concept of autocorrelation functions.

3.1 Autocorrelation Function and Time

The autocovariance function Γ provides a measure for the level of covariance between two samples in the chain of observables separated by t steps. Specifically, the functional value

$\Gamma(t)$ is obtained by averaging over the covariance of all such pairs and hence given by

$$\Gamma(t) = \frac{1}{M-t} \sum_{i=1}^{M-t} \text{Cov}(O_i, O_{i+t}). \quad (33)$$

It is conventional to normalise this to $t = 0$, turning the covariance into a correlation and thus yielding the autocorrelation function (ACF):

$$\rho(t) = \frac{\Gamma(t)}{\Gamma(0)}. \quad (34)$$

The IAT is defined as

$$\tau = 1 + 2 \sum_{t=1}^{\infty} \rho(t). \quad (35)$$

For Markov chains, the ACF is expected to follow an exponential decay such that for a large value of t , $\rho(t)$ will be dominated by noise. This motivates us to terminate the sum in eq. 35 at some separation K , to avoid inflating the inferred value of τ from noise contributions only. Selecting the value of the cut is discussed in [21] with the conclusion of the now conventional choice of $K = 4\tau$. Terminating the sum introduces an error in the estimate of τ whose closed form approximation was first introduced in [22] with a derivation sketched in [23]. In summary, we have

$$\tau \approx 1 + 2 \sum_{t=1}^K \rho(t), \quad K = 4\tau, \quad \text{Var}(\tau) \approx \frac{2(2K+1)}{M} \tau. \quad (36)$$

Finding K in practice requires terminating the sum at increasingly large separations and finding the first occasion when the condition $K \geq 4\tau$ is satisfied.

3.2 Efficient Computation

As evident from the above description, computing the ACF is numerically expensive and scales poorly with the number of observations. However, the simple observation that a covariance can be written as a convolution allows for an efficient computation using Fast Fourier transforms (FFTs). In more detail, denoting the discrete convolution by $*$, eq. 33 may be written as

$$\Gamma(t) \sim \sum_i (O_i - \bar{O})(O_{i+t} - \bar{O}) = \sum_i f_i f_{i+t} = f * f, \quad f_i \equiv O_i - \bar{O} \quad (37)$$

The cross-convolution theorem (proven in appendix A.2) states that a convolution in real space is equivalent to multiplication in Fourier space [24]. Hence, with \mathcal{F} denoting the Fourier transform, one concludes

$$\Gamma(t) \sim \mathcal{F}^{-1} [(\mathcal{F}[f])^2]. \quad (38)$$

Using this technique will be referred to as the FFT approach henceforth. In fact, this result applies more generally to the correlation function of two, potentially complex, scalar functions and will prove extremely useful when studying the $SU(2) \otimes SU(2)$ model, where a large number of correlations are frequently computed. In that context, we provide a run time comparison of the standard summation and FFT approach against the number of samples to correlate, showing that the latter method is faster by several orders of magnitudes. These results are presented in fig. 12. We further confirmed that the cumulative difference in the correlation function between the two approaches is of the order of 2% and likely noise-induced.

Care must be taken when using FFTs as their product yields the circular convolution of the original signals rather than the linear one. For physical correlations, the former automatically accounts for the periodic lattice boundary conditions and is therefore desirable. On the other hand, for autocorrelations in the chain of measurements, a linear convolution is required. An in-depth discussion of this can be found in [20], while here we will only highlight how this issue can be circumvented. From eq. 37 it becomes apparent that for a set $\{f_i\}$ of size M , the index $i+t$ can exceed M depending on the values of i, t . Supposing that the observations are periodic, meaning $f_j = f_{j \bmod M}$, causes no issue and results in a circular convolution. FFTs assume the observations to be of this type. In order to obtain a linear convolution, one must pad the data with sufficiently many zeros such that the circular mixing with these leaves the original data unchanged [25]. The padding is usually done in powers of 2 as FFT routines work most efficiently with such a number of elements. Once all Fourier transforms have been taken, one must truncate the introduced pads to get the final autocovariance function result.

4 Fourier Acceleration

4.1 Critical Slowing Down

We are now in a position to address the algorithmic motivation of this study: The power law increase of the simulation cost as the continuum is approached. This is commonly referred to as critical slowing down and is caused by the different evolution speeds of a system's Fourier modes. In particular, modes that evolve quickly set a lower bound for the integration step size ϵ below which the integrator becomes unstable and the trajectory significantly violates the conservation of the Hamiltonian, yielding a vanishing acceptance rate. This constraint causes the slow, long wavelength modes of the system to hardly evolve over the simulated fictitious time, resulting in large autocorrelations. In general, one seeks to extract macroscopic dynamics from lattice simulations which are contained in the slow modes and require simulations close to the continuum. It can be shown that the continuum limit of a discrete theory represents a critical point [26]. If the associated phase transition is second order or higher, the above-described difference in evolution speed intensifies and all correlation lengths and correlation times diverge. In particular, the scaling of the autocorrelation time τ with the correlation length ξ or the lattice spacing a is parameterised through the dynamical critical exponent z :

$$\tau \sim \xi^z \sim a^{-z}. \quad (39)$$

4.2 Modified Dynamics

Fourier Acceleration refers to a technique which aims to reduce (and for simple models even fully eliminate) the degree of critical slowing down by modifying the system's dynamics to force a more equal evolution speed. To illustrate how this is achieved, consider a free scalar field theory such that the HMC Hamiltonian is

$$H(\phi, \pi) = \frac{1}{2} \pi \cdot \pi + S(\phi), \quad S(\phi) = \frac{1}{2} \phi \cdot (m^2 - \partial^2) \cdot \phi = \frac{1}{2} \phi \cdot K(m) \cdot \phi. \quad (40)$$

The action kernel is denoted by $K(m)$ and for improved readability, details such as the summation over lattice sites have been absorbed into the “dot” notation as introduced in [27]. Hamilton's equations thus read

$$\dot{\phi} = \pi, \quad \dot{\pi} = -\frac{\partial S}{\partial \phi}. \quad (41)$$

The first equation should be understood as introducing the conjugate momenta as the time derivative of the degrees of freedom of the theory. Combining Hamilton's equations and seeking plane wave solutions for ϕ leads to the dispersion relation

$$\omega^2(p) = m^2 + p^2 \quad (42)$$

which reveals the evolution speed dependence on the mode's momentum. The finite lattice volume and spacing limit respectively introduce a lower and upper bound on the momentum and thus limit the range of possible evolution frequencies. The ratio

$$\frac{\omega_{max}^2}{\omega_{min}^2} = \frac{m^2 + p_{max}^2}{m^2 + p_{min}^2} \quad (43)$$

therefore characterises the amount of critical slowing down and diverges as the lattice is taken to the continuum, causing $p_{max} \rightarrow \infty, p_{min} \rightarrow 0$.

Inserting the inverse of the action kernel in the kinetic energy allows to modify this ratio. Namely,

$$H(\phi, \pi) = \frac{1}{2} \pi \cdot K(M)^{-1} \cdot \pi + S(\phi) \implies \omega^2(p) = \frac{m^2 + p^2}{M^2 + p^2} \quad (44)$$

where the acceleration mass M is introduced as an additional model parameter. Changing the value of M affects the dispersion relation and consequently the degree of critical slowing down. For the free theory considered here, choosing M to equal the physical mass removes the momentum dependence in ω and thus fully avoids critical slowing down. We show this explicitly in the study of the quantum harmonic oscillator in sec. 5.2. For more complicated systems, such as the $SU(2) \otimes SU(2)$ model, obtaining perfect acceleration is more difficult as the choice of M is non-trivial. Finding the optimal value of M requires a metric measuring the efficiency of the simulation. To that end, we construct a cost function of the following form:

$$\text{cost}(M) = \frac{\text{simulation time}}{\text{acceptance rate}} \sqrt{\tau}. \quad (45)$$

This is motivated by the measurement error scaling as $\sqrt{\tau}$ and efficient simulations are desired to run quickly with a high acceptance rate. Otherwise, resources are wasted on the computation of frequently rejected candidate configurations. Since the units of the cost function are arbitrary, only ratios are meaningful.

Modifying the Hamiltonian comes at the price of increasing the complexity of the ϕ equation of motion. Specifically, Hamilton's equations based on eq. 44 now read

$$\dot{\phi} = K(M)^{-1} \cdot \pi, \quad \dot{\pi} = -\frac{\partial S}{\partial \phi}. \quad (46)$$

Comparing with the relation between position, momentum, and mass in classical dynamics allows to view the modification of the kinetic energy as introducing an artificial mass $K(M)$ whose value is momentum dependent. In particular, the evolution of fast modes is slowed by getting assigned a large mass value while slow modes are accelerated through small masses, resulting in overall more uniform evolution rates [28]. Computing K^{-1} in configurational space is possible but the matrix inversion becomes expensive for large lattice volumes. However, in Fourier space $K(M)$ is diagonal such that the inversion becomes trivial. The ϕ update is therefore performed in Fourier space, giving rise to the name Fourier Acceleration. How to do this in practice will be the topic of the following section.

4.3 Acceleration in Practice

To illustrate the details of Fourier Acceleration (FA), we will continue to use the free scalar field example from the previous section. This framework will then be used in later sections to study the one-dimensional quantum harmonic oscillator and the $SU(2) \otimes SU(2)$ model in two dimensions. We will begin with the description of acceleration in one dimension and later illustrate the generalisation to higher dimensions. The following discussion is mainly based on [27].

We adopt the following convention for discrete Fourier transforms:

$$\mathcal{F}^{-1}(\tilde{\phi}) \equiv \phi_n = \frac{1}{L} \sum_{k=0}^{L-1} \tilde{\phi}_k e^{2\pi i \frac{kn}{L}} \quad (47)$$

where n, k index the real and Fourier space lattice sites respectively i.e. $n, k = 0, \dots, L-1$. Reintroducing the lattice sums in eq. 40 and discretizing the second derivatives yields

$$S(\phi) = \frac{1}{2} \sum_{n,m} \phi_n K_{nm} \phi_m, \quad K_{nm} = m^2 \delta_{mn} - (\delta_{n+1,m} - 2\delta_{nm} + \delta_{n-1,m}). \quad (48)$$

Upon inserting the Fourier representation of the field, it follows that

$$S(\phi) = \frac{1}{2L^2} \sum_{n,m} \sum_{k,k'} \tilde{\phi}_k K_{nm} \tilde{\phi}_{k'} e^{2\pi i \frac{kn+k'm}{L}} \quad (49)$$

$$= \frac{1}{2L^2} \sum_{n,m} \sum_{k,k'} \tilde{\phi}_k \left(2 - e^{2\pi i \frac{k'}{L}} - e^{-2\pi i \frac{k'}{L}} + m^2 \right) \tilde{\phi}_{k'} e^{2\pi i \frac{n(k+k')}{L}}. \quad (50)$$

Recalling a basic trigonometric identity allows to combine the first 3 terms, while the sum over n together with the remaining exponential factor can be identified as the Fourier representation of the Dirac delta. Hence

$$S(\phi) = \frac{1}{2L^2} \sum_m \sum_{k,k'} \tilde{\phi}_k \left(4 \sin^2 \frac{\pi k}{L} + m^2 \right) \tilde{\phi}_{k'} \delta_{k,-k'} \quad (51)$$

$$= \frac{1}{2L} \sum_{k,k'} \tilde{\phi}_k \left(4 \sin^2 \frac{\pi k}{L} + m^2 \right) \tilde{\phi}_{k'} \delta_{k,-k'}. \quad (52)$$

Note that due to the periodic boundary conditions, the lattice site at $-k'$ is identified with the site at $L - k'$. In this report, the degrees of freedom of the considered systems will always be real, such that their Fourier transforms are required to satisfy Hermitian symmetry: $\tilde{\phi}_{L-k} = \tilde{\phi}_k^*$. We therefore obtain

$$S(\phi) = \frac{1}{2L} \sum_{k,k'} \tilde{\phi}_k \left(4 \sin^2 \frac{\pi k}{L} + m^2 \right) \tilde{\phi}_{k'} \delta_{k,L-k'} \quad (53)$$

$$= \frac{1}{2L} \sum_k \tilde{\phi}_k^* \left(4 \sin^2 \frac{\pi k}{L} + m^2 \right) \tilde{\phi}_k = \frac{1}{2L} \tilde{\phi}^\dagger \cdot \tilde{K}(m) \cdot \tilde{\phi}. \quad (54)$$

Crucially, the Fourier space kernel is diagonal with inverse

$$A_k \equiv \tilde{K}_k^{-1}(M) = \frac{1}{4 \sin^2 \frac{\pi k}{L} + M^2}. \quad (55)$$

In fact, the above process does not change much when generalising to a D -dimensional lattice. Rather than indexing lattice size by a single value k , a vector $\mathbf{k} = \{k_1, \dots, k_D\}$ is required. Assuming a square lattice, each component will range between 0 and $L - 1$ and summing over the vector can be understood as summing over all its components. In particular, noting the normalisation of the Fourier transform changes to $1/L^D$, we have

$$S(\phi) = \frac{1}{2L^{2D}} \sum_{\mathbf{n}, \mathbf{m}} \sum_{\mathbf{k}, \mathbf{k}'} \tilde{\phi}_{\mathbf{k}} K_{\mathbf{n}\mathbf{m}} \tilde{\phi}_{\mathbf{k}'} e^{2\pi i \frac{\mathbf{k} \cdot \mathbf{n} + \mathbf{k}' \cdot \mathbf{m}}{L}}. \quad (56)$$

Repeating the steps of the above discussion for each component of \mathbf{k} yields

$$A_{\mathbf{k}} = A_{k_1, \dots, k_d} = \tilde{K}_{\mathbf{k}}^{-1}(M) = \frac{1}{4 \sin^2 \frac{\pi k_1}{L} + \dots + 4 \sin^2 \frac{\pi k_D}{L} + M^2}. \quad (57)$$

It is convenient to think of $A_{\mathbf{k}}$ as a (symmetric) field defined on the Fourier space lattice which modifies the ϕ equation of motion. In particular, turning the modified field equation of motion of eq. 46 into Fourier space gives

$$\dot{\tilde{\phi}}_{\mathbf{k}} = A_{\mathbf{k}} \tilde{\pi}_{\mathbf{k}} \quad (58)$$

with no sum over \mathbf{k} . Obtaining the desired real space evolution is now just a matter of taking the inverse Fourier transform:

$$\dot{\phi}_{\mathbf{n}} = \mathcal{F}^{-1} [A_{\mathbf{k}} \tilde{\pi}_{\mathbf{k}}]. \quad (59)$$

For later reference, one may also write the Hamiltonian immediately in Fourier space to deduce the associated equations of motion:

$$H(\phi, \pi) = \tilde{H}(\tilde{\phi}, \tilde{\pi}) = \frac{1}{2} \tilde{\pi}^\dagger \cdot \tilde{K}^{-1}(M) \cdot \tilde{\pi} + \tilde{S}(\tilde{\phi}). \quad (60)$$

4.4 Distribution of Momenta

By modifying the kinetic term, the momentum distribution becomes more complicated and is again easiest sampled from in Fourier space. Based on eq. 60, we have

$$\tilde{\pi} \sim \exp\left(-\frac{1}{2}\tilde{\pi}^\dagger \cdot \tilde{K}^{-1} \cdot \tilde{\pi}\right) = \exp\left(-\frac{1}{2L^D} \sum_{\mathbf{k}} |\tilde{\pi}_{\mathbf{k}}|^2 A_{\mathbf{k}}\right). \quad (61)$$

Sampling from this distribution and taking an inverse Fourier transform after yields the momentum in real space, required for the leapfrog integration. Again, by requiring π to be real, Hermitian symmetry is imposed on the Fourier modes:

$$\tilde{\pi}_{k_1, \dots, k_D} = \tilde{\pi}_{L-k_1, \dots, L-k_D}^*. \quad (62)$$

As only one- and two-dimensional lattices are considered in this report, we will show explicitly how this constraint can be implemented for such systems. The generalisation to higher dimensions will become apparent from these examples and is further discussed in [27]. Throughout, a cubic lattice of even length L is assumed. For notational convenience, we denote real and imaginary components by superscripts, such that

$$\tilde{\pi}_{\mathbf{k}} = \tilde{\pi}_{\mathbf{k}}^R + i\tilde{\pi}_{\mathbf{k}}^I. \quad (63)$$

One dimension Recall eq. 47 and note that the exponential factor $\exp(-2\pi i \frac{kn}{L})$ becomes real for $k = 0, \frac{L}{2}$, such that $\tilde{\pi}_k$ must be real in order to produce real valued π_n . In other words,

$$\tilde{\pi}_0^I = \tilde{\pi}_{\frac{L}{2}}^I = 0 \quad (64)$$

while the real components are unconstrained. Due to hermiticity, the remaining momentum components are fully determined through $\tilde{\pi}_k^R$ and $\tilde{\pi}_k^I$ for $k = 1, \dots, \frac{L}{2} - 1$ since

$$\tilde{\pi}_{L-k} = \tilde{\pi}_k^R - i\tilde{\pi}_k^I. \quad (65)$$

Overall, only L of the total $2L$ real and imaginary components are independent and a particularly convenient way of organising these is to define

$$\tilde{\Pi} \equiv \left(\tilde{\pi}_0^R, \sqrt{2}\tilde{\pi}_1^R, \dots, \sqrt{2}\tilde{\pi}_{\frac{L}{2}-1}^R, \tilde{\pi}_{\frac{L}{2}}^R, \sqrt{2}\tilde{\pi}_{\frac{L}{2}-1}^I, \dots, \sqrt{2}\tilde{\pi}_1^I\right). \quad (66)$$

This object is real and its components are denoted as $\tilde{\Pi}_k, k = 0, \dots, L-1$. Further observe that

$$|\tilde{\pi}_0|^2 = |\tilde{\pi}_0^R|^2 \quad \text{and} \quad \sum_{k=1}^{\frac{L}{2}-1} |\tilde{\pi}_k|^2 + \sum_{k=\frac{L}{2}}^{L-1} |\tilde{\pi}_k|^2 = 2 \sum_{k=1}^{\frac{L}{2}-1} |\tilde{\pi}_k|^2, \quad (67)$$

such that $\sum_k |\tilde{\pi}_k|^2 = \sum_k \tilde{\Pi}_k^2$, allowing to express the one-dimensional version of eq. 61 as

$$\tilde{\pi} \sim \exp\left(-\frac{1}{2L} \sum_{k=0}^{L-1} \tilde{\Pi}_k^2 A_k\right) = \prod_{k=0}^{L-1} \exp\left(-\frac{1}{2L} \tilde{\Pi}_k^2 A_k\right). \quad (68)$$

The components $\tilde{\Pi}_k$ can now be sampled from independent Gaussians with zero mean and standard deviation $\sqrt{L/A_k}$. Using eq. 66 together with eqs. 64, 65 allows to fully reconstruct $\tilde{\pi}_k$ from $\tilde{\Pi}_k$. Taking the Fourier transform will yield real-valued π_n that follows the modified momentum distribution of eq. 44.

Two dimensions In this case $\mathbf{k} = (k_1, k_2)$, and can be understood as repeating the one-dimensional process on the first index, while keeping the second fixed. Doing so for all values of k_2 completes the process, yielding L^2 independent components which are organised into $\tilde{\Pi}_{k_1, k_2}$ where each row follows the structure of eq. 66. In fact, these steps can be combined in the following way.

The special case that the Fourier mode must be real only occurs when both exponential factors in the transform are real, i.e. when $k_1, k_2 \in \{0, \frac{L}{2}\}$. Thus,

$$\tilde{\pi}_{0,0}^I = \tilde{\pi}_{\frac{L}{2},0}^I = \tilde{\pi}_{0,\frac{L}{2}}^I = \tilde{\pi}_{\frac{L}{2},\frac{L}{2}}^I = 0 \quad (69)$$

while their associated real components are unconstrained. If just one of k_1, k_2 takes either of these special values, one exponential factor remains complex such that the only constraint comes from the Hermitian symmetry. For $k = 1, \dots, \frac{L}{2} - 1$, this implies

$$\tilde{\pi}_{L-k,0} = \tilde{\pi}_{k,0}^*, \quad \tilde{\pi}_{0,L-k} = \tilde{\pi}_{0,k}^*, \quad \tilde{\pi}_{L-k,\frac{L}{2}} = \tilde{\pi}_{k,\frac{L}{2}}^*, \quad \tilde{\pi}_{\frac{L}{2},L-k} = \tilde{\pi}_{\frac{L}{2},k}^*. \quad (70)$$

For all other cases, namely $k_1, k_2 = 1, \dots, \frac{L}{2} - 1$, hermiticity requires

$$\tilde{\pi}_{k_1,k_2} = \tilde{\pi}_{L-k_1,L-k_2}^* \quad \text{and} \quad \tilde{\pi}_{L-k_1,k_2} = \tilde{\pi}_{k_1,L-k_2}^*. \quad (71)$$

Again, $\tilde{\Pi}_{k_1,k_2}$ is real and constructed such that $\sum_{k_1,k_2} |\tilde{\pi}_{k_1,k_2}|^2 = \sum_{k_1,k_2} \tilde{\Pi}_{k_1,k_2}^2$. Hence

$$\tilde{\pi} \sim \prod_{k_1,k_2} \exp\left(-\frac{1}{2L^2} \tilde{\Pi}_{k_1,k_2}^2 A_{k_1,k_2}\right), \quad (72)$$

where $\tilde{\Pi}_{k_1,k_2}$ are sampled from independent Gaussians with zero mean and standard deviation $\sqrt{L^2/A_{k_1,k_2}}$. Reconstructing $\tilde{\pi}$ from $\tilde{\Pi}$ and taking the inverse Fourier transform yields the required real space momentum samples.

Generalising to higher dimensions requires us to observe that for every additional dimension introduced, the process reduces to repeating the steps from the preceding dimension and is therefore a recursive application of the one-dimensional case.

The equipartition theorem offers a straightforward way to test if this sampling procedure is implemented correctly. On average each quadratic degree of freedom³ contributes a factor of $1/2$ to the total kinetic energy in the chosen convention of $k_B = T = 1$. This is to be compared against the average modified kinetic energy computed based on generating many momentum samples following the above process.

³In a scalar theory this is 1 per lattice site, while for the $SU(2) \otimes SU(2)$ model it is 3.

4.5 The Fourier Accelerated Hamiltonian Monte Carlo Algorithm

It should be noted that apart from the modification of the equations of motion and the more complicated sampling of the auxiliary momenta, Fourier Acceleration integrates very naturally with HMC. The resulting Fourier accelerated Hamiltonian Monte Carlo (FA HMC) algorithm shows all the benefits of standard HMC in addition to significantly reducing autocorrelations. Algorithm 2 gives the FA HMC pseudocode. Further, as described in [28], Fourier Acceleration can also be applied to gauge theories, making FA HMC highly attractive for QCD lattice simulations [6, 29].

Algorithm 2 Fourier accelerated Hamiltonian Monte Carlo

Input: Action $S(\phi)$ and the inverse kernel, such that $H(\phi, \pi) = \frac{1}{2}\pi \cdot K^{-1}(M) \cdot \pi + S(\phi)$.
Acceleration mass parameter M , lattice length L and dimensions D .

Leapfrog parameters: step size ϵ , number of steps ℓ .

Length of Markov chain \mathcal{M} .

Output: Markov chain $\{\phi_i\} = \{\phi_1, \dots, \phi_{\mathcal{M}}\}$ with configurations following $P(\phi) = e^{-S(\phi)}$ once the chain has equilibrated.

- 1: Provide a starting configuration ϕ_0
 - 2: Construct $A_{\mathbf{k}} = \tilde{K}_{\mathbf{k}}^{-1}(M)$
 - 3: **for** $i = 1$ to \mathcal{M} **do**
 - 4: Draw a momentum sample π :
 - 5: Sample $\tilde{\Pi}_{\mathbf{k}} \sim \text{Normal}(0, \sqrt{L^D/A_{\mathbf{k}}})$
 - 6: Reconstruct $\tilde{\pi}_{\mathbf{k}}$ from $\tilde{\Pi}_{\mathbf{k}}$
 - 7: Take inverse Fourier transform $\pi = \mathcal{F}^{-1}[\tilde{\pi}_{\mathbf{k}}]$
 - 8: Evolve Hamilton's equations for time $\epsilon\ell$: $(\phi, \pi) \rightarrow (\phi^*, \pi^*)$ via Leapfrog($\phi, \pi, \ell, \epsilon$):
 - 9: Update momentum: $\dot{\pi} = -\frac{\partial S}{\partial \phi}$
 - 10: Update field: $\dot{\phi} = \mathcal{F}^{-1}[A_{\mathbf{k}}\mathcal{F}[\pi]]$
 - 11: Reverse the candidate momentum: $\pi^* \rightarrow -\pi^*$
 - 12: Compute the acceptance probability: $\alpha = \min(1, e^{-(H(\phi^*, \pi^*) - H(\phi, \pi))})$
 - 13: Draw a uniform random number: $u \sim \text{Uniform}(0, 1)$
 - 14: **if** $\alpha > u$ **then**
 - 15: Accept the candidate as the new configuration: $\phi_{i+1} = \phi^*$
 - 16: **else**
 - 17: Reject the candidate and repeat the current configuration: $\phi_{i+1} = \phi$
 - 18: **end if**
 - 19: **end for**
-

Having introduced the required machinery, we now apply Fourier Acceleration to two physical systems with the aim to quantify the improvements FA HMC offers compared to standard HMC close to the continuum limit.

5 Quantum Harmonic Oscillator

Before studying the $SU(2) \otimes SU(2)$ model, we use the quantum harmonic oscillator (QHO) as a testbed for our simulation implementation. As this system is analytically solvable both on a lattice and in the continuum limit, we are able to validate the simulation results and rule out any systematic errors which might have been difficult to spot when simulating $SU(2) \otimes SU(2)$ dynamics. We will consequently keep this section brief and use this simple model to introduce techniques which will be used in the later $SU(2) \otimes SU(2)$ study. The continuum and lattice predictions of the QHO as well as some additional discussions can be found in Appendix B.

In particular, we consider a one-dimensional quantum harmonic oscillator of mass m on a lattice of length L with spacing a . Lattice sites are indexed through $i = 0, 1, \dots, L-1$ and periodic boundary conditions are imposed, implying the identification $i = i \bmod L$. To make the general discussion of the previous sections more concrete, we change the notation from (ϕ, π) to (x, p) where the position and momentum are both defined on the lattice and should be therefore thought of periodic vectors with L components. A lattice configuration is therefore given by a particular realisation of $x = (x_0, \dots, x_{L-1})$. Discretizing the standard QHO action and introducing a simple kinetic energy quadratic in the momenta yields the Hamiltonian

$$H(x, p) = T(p) + S(x) = \sum_{i=0}^{L-1} \frac{p_i^2}{2} + \sum_{i=0}^{L-1} a \left[\frac{m}{2} \left(\frac{x_{i+1} - x_i}{a} \right)^2 + \frac{1}{2} m \omega^2 x_i^2 \right]. \quad (73)$$

Noting that two terms in the second sum contain x_i allows us to quickly find

$$\frac{\partial S}{\partial x_i} = \frac{1}{a} [m(2x_i - x_{i-1} - x_{i+1}) + a^2 m \omega^2 x_i], \quad \frac{\partial T}{\partial p_i} = p_i. \quad (74)$$

With these results, the HMC logic, and the leapfrog scheme as presented in sec. 2.4, lattice configurations are generated with probability weight $\exp(-H(x, p))$. These are then used to measure various observables with the results contrasted to analytic predictions.

5.1 Measurement Results

A well-known property of the QHO is the non-trivial expectation value of the squared position, which will be the first measurement of the simulations. For each configuration in the chain, the average value of x^2 will be computed before using this data to find the ensemble average and associated statistical error. Initialising the chain with a random configuration and running the simulation on a $L = 100, a = 1$ lattice for an oscillator of unit mass and frequency produces the chain of measurements displayed in fig. 2.

The analytic lattice prediction is shown by the solid red line with the explicit expression given in appendix B. The good agreement between this prediction and the numerical results after only 1000 trajectories is due to the simplicity of the system and the low dimensionality.

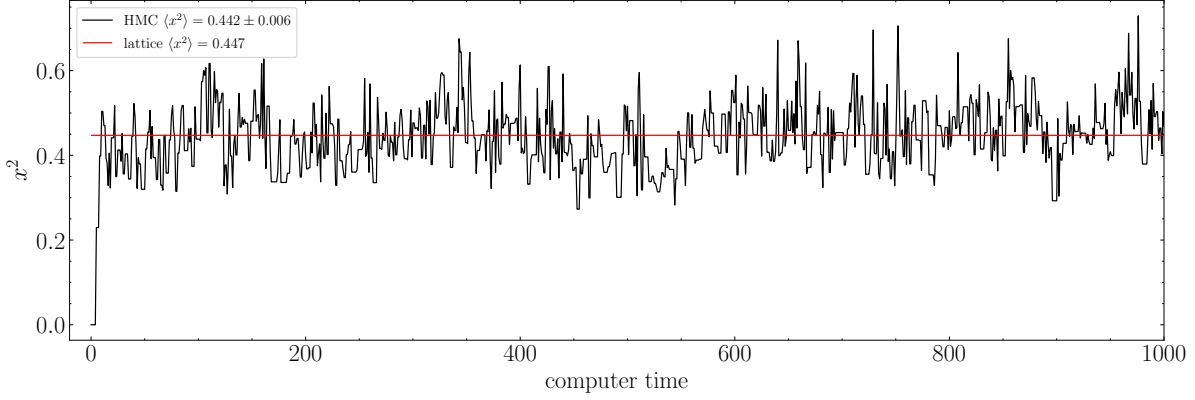


Figure 2. Measurement chain for the squared position of a simulation using $L = 100, a = 1, m = 1, \omega = 1$ with no burn in rejected to illustrate the fast thermalisation. The analytically expected value of $\langle x^2 \rangle$, as well as the measurement with the associated uncertainty, are presented in the legend.

These factors further allow the chain to converge to the typical set very quickly, as seen by the short thermalisation process of $\mathcal{O}(10)$ trajectories. A further feature illustrated by the plot is the short IAT visible in the lack of structure in the chain of measurements. For the presented run, $\tau = 12.20 \pm 3.84$ was computed, where the large error is due to the small number of trajectories. Comparing fig. 2 to plots in the literature that use the Metropolis algorithm instead of HMC (such as [30, 31]) underlines the superior efficiency of HMC, as in the pure Metropolis case autocorrelations are greater and the chain converges more slowly. Note that since a finite lattice spacing is used, the lattice prediction differs from the commonly remembered continuum result $\langle x^2 \rangle = 0.5$ for $m = \omega = 1$.

There exist a variety of more complex checks to validate that the simulation results agree with the lattice theory prediction such as computing the ground state energy and wave function. As these discussions are not directly relevant to the further progress of the project, they are omitted here but presented in appendix B.2 for the interested reader. Instead, we focus on estimating the energy gap between the ground state and the first excited state at different values of the lattice spacing. This will illustrate the convergence of lattice methods (numerical and analytic) to the continuum result while also introducing the concept of an effective mass which will also be used in the $SU(2) \otimes SU(2)$ study.

In general, the correlation function between two lattice sites is expected to decay exponentially with their physical separation. For a quantum mechanical system with a vanishing position expectation value, the rate of this decay, known as the effective mass, is closely related to the lowest energy gap. The two-point correlation function is thus

$$C(d) = \langle \hat{x}(0)\hat{x}(d) \rangle - \langle \hat{x}(0) \rangle \langle \hat{x}(d) \rangle \quad (75)$$

$$= \langle 0 | \hat{x}(0)\hat{x}(t) | 0 \rangle \quad (76)$$

and following [9] yields

$$C(d) = \sum_{n \neq 0} |\langle 0 | \hat{x} | n \rangle|^2 e^{-(E_n - E_0)d} \approx |\langle 0 | \hat{x} | 1 \rangle|^2 e^{-(E_1 - E_0)d}. \quad (77)$$

where $\hbar = 1$. For the first equality, a complete set of energy eigenstates $|n\rangle$ was inserted between the position operators and the relation between Heisenberg and Schrödinger operators was used. Supposing E_n increases with n and considering the large d limit allows to conclude that the dominant contribution comes from the $n = 1$ term and thus yields the final expression.

Analytic correlation function In the above derivation, the periodic boundary conditions are not accounted for. These cause lattice points separated by d and $L - d$ lattice spacings to have the same physical distance between them. One, therefore, expects the correlation function to follow the shape of a cosh rather than an exponential decay. Up to a proportionality constant, which can be removed by normalising to $d = 0$, the analytically expected form of the correlation function is

$$C(d) \sim \cosh\left(\frac{d - L/2}{\xi}\right) - 1. \quad (78)$$

Fitting normalised numerical data to eq. 78 allows to extract the correlation length ξ in units of the lattice spacing. One can exploit the symmetry of the correlation function induced by the periodic boundary conditions to effectively increase the statistics by a factor of 2 through averaging the data at d and $L/2 + d$ for $d \in [0, L/2]$. The uncertainty in the averaged value will be reduced by a factor of $1/\sqrt{2}$ compared to the original error. We will refer to this technique as mirroring the correlation function. Apart from this statistical error, which can be controlled by longer simulations, choosing the range to fit over introduces a systematic uncertainty. Often manual specification of the range based on examining the correlation function plot is the only accessible solution. This will be the case for the $SU(2) \otimes SU(2)$ model.

However, since the QHO is analytically solvable, we use the lattice theory prediction of the correlation length from [9] to define the fitting range. Noting that ξ is measured in units of the lattice spacing implies that the energy gap is given by $1/(\xi a)$. In fig. 3, the inferred energy gap is plotted against the value of the lattice spacing for fixed $L = 100, m = 1, \omega = 1$. Each data point is the result of performing the above-described fitting procedure to the correlation function which was obtained from simulating 5×10^5 trajectories. Since the thermalisation period increases as $a \rightarrow 0$, we rejected the initial 10% of the trajectories for all runs. The analytic predictions from lattice and continuum theory, based on eq. 159, are shown as solid lines in fig. 3. We concentrated our computation on the small a region where we expect the analytic and numerical results to converge.

Two main conclusions can be drawn from this figure. When excluding the first data point, the numerical results agree with the analytic lattice prediction at the sub-percent level for the range of considered lattice spacings. The growing error bars as $a \rightarrow 0$ can be understood in the context of critical slowing down as increasing autocorrelation times, causing greater variance in the correlation function measurements. The second conclusion is that lattice computations can yield continuum results only in a narrow range of lattice spacings. While the continuum is poorly approximated for large lattice spacings, considering small values of a

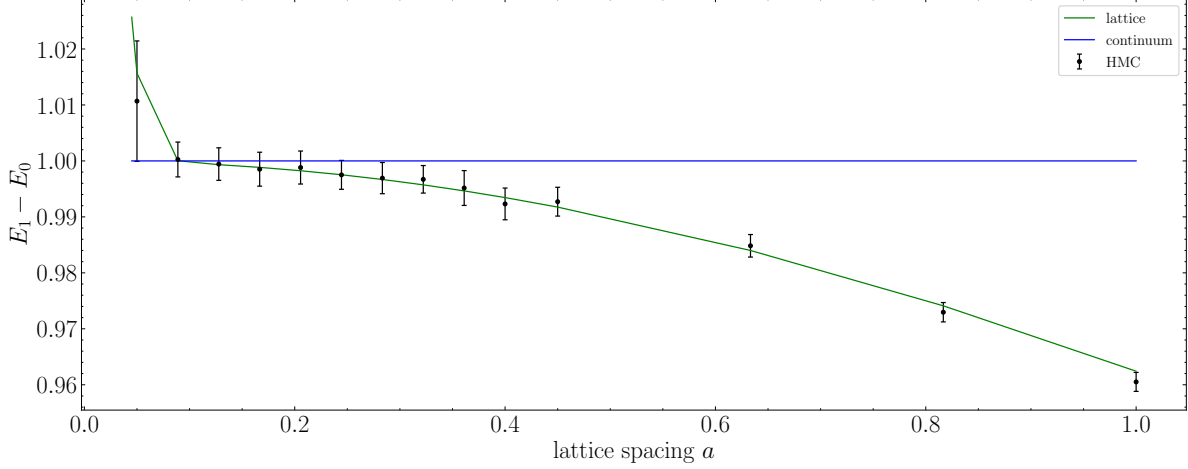


Figure 3. The energy gap between the ground state and first excited state against the lattice spacing for $L = 100, m = 1, \omega = 1$. For each data point, 5×10^5 trajectories were simulated with a 10% burn in rejection. At the lower end of the survey, larger autocorrelation times yield increasing error bars and finite size effects cause the lattice approach to break down, visible in the divergence from the continuum result. It becomes apparent that lattice predictions only yield close to continuum results in a narrow range of a for fixed L .

at a fixed lattice size reduces the physical size of the lattice and causes finite size effects. This is the origin for the observed divergence of the lattice results from the continuum prediction as $a \rightarrow 0$.

5.2 Fourier Acceleration

In this section, we show that adding FA to the QHO simulations allows to fully avoid critical slowing down. To that end, we determine the critical exponent z of eq. 39 for standard HMC and for FA HMC by computing the position IAT from simulations at different values of the lattice spacing. Fitting a power law to the data allows to extract z .

First, we describe how Fourier Acceleration is implemented for the QHO, using the framework of sec. 4.3. To begin, one needs to identify the action kernel:

$$S(x) = \sum_i a \left(\frac{m}{2} \left(\frac{x_{i+1} - x_i}{a} \right)^2 + \frac{1}{2} m \omega^2 x_i^2 \right) \quad (79)$$

$$= \frac{1}{2} \sum_i \frac{m}{a} (x_{i+1}^2 - 2x_{i+1}x_i + x_i^2) + am\omega^2 x_i^2 \quad (80)$$

$$= \frac{1}{2} \sum_{i,j} x_i K_{ij} x_j \quad (81)$$

$$\implies K_{ij} = \frac{m}{a} (2\delta_{i,j} - \delta_{i-1,j} - \delta_{i+1,j} + a^2 \omega^2 \delta_{i,j}) . \quad (82)$$

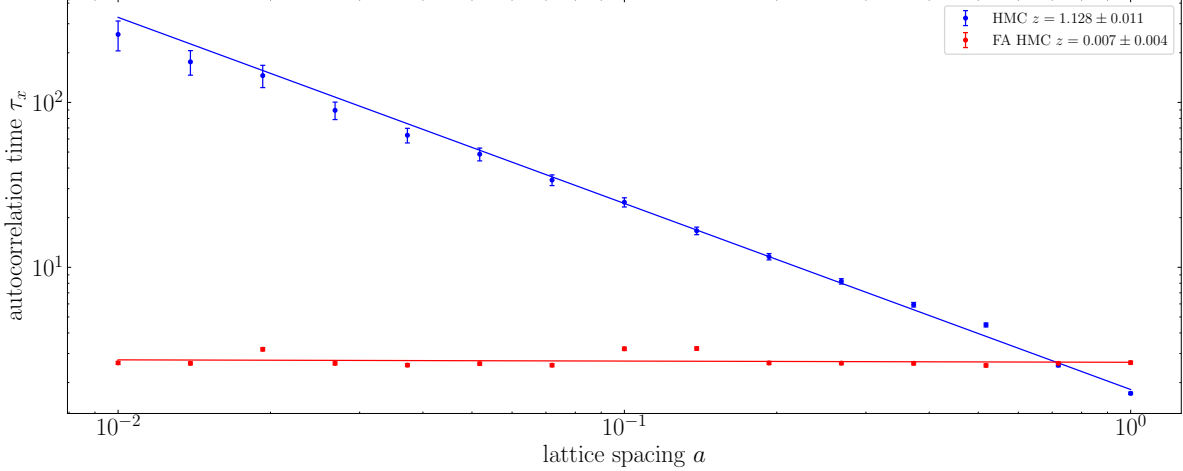


Figure 4. Numerical data of the position autocorrelation time τ_x of the quantum harmonic oscillator against the value of the lattice spacing a on log-log axes for standard HMC (blue) and FA HMC (red). Fitted power laws are superimposed with the value of the critical exponent presented in the legend. Using FA HMC allows to avoid critical slowing down entirely. Each data point corresponds to 10^6 (10% burn in rejection) trajectories on a $L = 1000$ lattice.

The last step exploits that due to periodic boundary conditions, we have $\sum_i x_i = \sum_i x_{i+1}$. The deep connection between free field theory and the harmonic oscillator is manifested in the identical structure of their kernels and thus allows to conclude

$$\tilde{K}_{k,k'} = \frac{m}{a} \left(4 \sin^2 \frac{\pi k}{L} + a^2 \omega^2 \right) \delta_{k,k'} \quad (83)$$

based on the result of sec. 4.3. Note that the above derivation dictates the choice for the acceleration mass parameter which is expected to yield perfect acceleration⁴. Denoting the kernel inverse by A_k and using eq. 60 results in the modified Hamiltonian

$$H(x, p) = \frac{1}{2L} \sum_k A_k |\tilde{p}_k|^2 + S(x). \quad (84)$$

The collected IAT data as well as the fitted power laws are shown in fig. 4. Each data point was obtained by simulating 10^6 trajectories (10% burn in rejection) for $L = 1000, m = 1$ and $\omega = 1$. Due to the considered small values of the lattice spacing, a large lattice size was necessary to avoid finite size effects. For FA HMC, z vanishes within statistical uncertainty, representing that critical slowing down is entirely avoided. For ordinary HMC, $z = 1.128 \pm 0.011$ was inferred, such that at $a = 0.01$ FA HMC allows to gain 2 orders of magnitude in the position IAT. Recall from eq. 32 that for M trajectories, the statistical error of a measurement scales with $\sqrt{\tau/M}$. To achieve precision comparable to that of FA HMC with standard HMC would therefore require 100 times as many trajectories to be simulated.

A more complete view of the efficiency improvement through FA is described by the cost function, defined in eq. 45. The ratio between the standard HMC and FA HMC cost is

⁴This property breaks down for more complicated systems, such as the anharmonic oscillator.

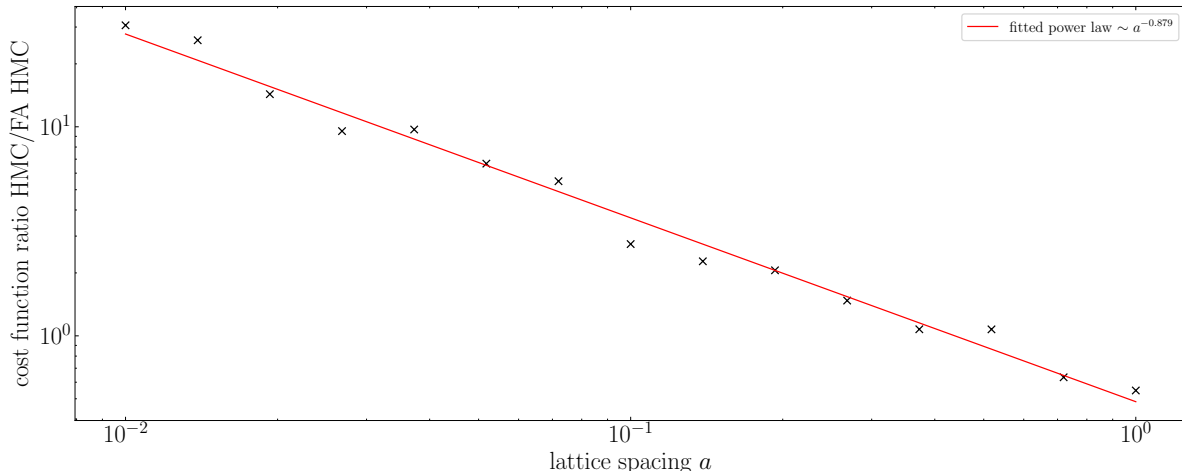


Figure 5. Ratio of the quantum harmonic oscillator cost function defined in eq. 45 for standard HMC and FA HMC against the lattice spacing a on a log-log scale with a fitted power law superimposed. The simulation parameters are identical to those used in fig. 4.

plotted in fig. 5 and appears to follow a power law with a fitted exponent of -0.879 , allowing us to quantify the improvement obtained through Fourier Acceleration as the continuum is approached. Note that far away from criticality, the additional computations in FA HMC cause its computational cost to grow larger than that of standard HMC such that the ratio falls below 1.

However, as the QHO is analytically solvable, resorting to numerical studies is not necessary and does not offer any additional insight. This is contrasted with the $SU(2) \otimes SU(2)$ model where simulations give access to key physics.

6 $SU(2) \otimes SU(2)$ Chiral Model in 2 Dimensions

6.1 Motivation

The study of chiral models is interesting in its own right, mainly due to the rich phase structure. However, it is also worth studying these models from a particle physics perspective as they share many similarities with non-abelian gauge theories while being computationally easier to work with. In both cases the degrees of freedom are Lie group elements, yielding similar mathematical features and challenges. Specifically, as first shown by Migdal [32], the renormalization group trajectories of $SU(N) \otimes SU(N)$ spin systems in D dimensions are identical to those of $SU(N)$ gauge theories in $2D$ dimensions. Further essential shared features are confinement and asymptotic freedom, with a more complete overview given in [33].

The $N = 2$ case in 2 dimensions is particularly interesting as it is known to lie in a single phase for all values of the coupling parameter [34]. This single confining and asymptotically

free phase captures the key physics of QCD while being computationally inexpensive to simulate. In this section, we describe how standard HMC and FA HMC simulations can be constructed for this model. Before examining asymptotic freedom numerically, we compute the internal energy density in a regime of the coupling parameter which is difficult to access analytically. Finally, the gain in simulation efficiency through FA HMC will be quantified.

6.2 $SU(2)$ Matrix Properties

To begin, some basic properties of $SU(2)$ matrices in the fundamental representation are introduced. The simulation code is constructed to act entirely on such matrices and by exploiting their properties the efficiency of the implementation is increased considerably.

We choose the Pauli matrices σ_i as the $SU(2)$ group generators and parameterise the group elements by $\boldsymbol{\alpha} \in \mathbb{R}^3$, representing the 3 degrees of freedom of an $SU(2)$ matrix. Using summation convention throughout, the simply connectedness of $SU(N)$, the explicit form of the Pauli matrices and their important property

$$\sigma_i \sigma_j = \delta_{ij} \mathbb{1}_2 + i \epsilon_{ijk} \sigma_k, \quad (85)$$

the following sequence of equalities can be easily verified:

$$SU(2) \ni U = e^{i\boldsymbol{\alpha} \cdot \boldsymbol{\sigma}} = \cos \alpha \mathbb{1}_2 + i \sin \alpha (\hat{\boldsymbol{\alpha}} \cdot \boldsymbol{\sigma}) \quad (86)$$

$$= a_0 \mathbb{1}_2 + i a_i \sigma_i \quad (87)$$

$$= \begin{pmatrix} a_0 + i a_3 & a_2 + i a_1 \\ -a_2 + i a_1 & a_0 - i a_3 \end{pmatrix}. \quad (88)$$

In the third equality, the parameter vector $a_\mu = (a_0, a_i)$ with $a_0 = \cos \alpha, a_i = \hat{\boldsymbol{\alpha}}_i \sin \alpha$ was introduced where $\alpha = |\boldsymbol{\alpha}|$ and $\hat{\boldsymbol{\alpha}} = \boldsymbol{\alpha}/\alpha$. The unitarity constraint and the $\det U = 1$ requirement yield that the norm of a_μ is unity:

$$|a|^2 = a_0^2 + a_1^2 + a_2^2 + a_3^2 = 1. \quad (89)$$

The 3-sphere S_3 can thus be identified as the group manifold. It is not only numerically convenient to represent U by a_μ , but it also allows to interpret an $SU(2)$ valued matrix field as a collection of 4 real scalar fields, which are subject to the above constraint. Further properties employed in the implementation are listed below.

Matrix product Consider $A, B \in SU(2)$ with respective parameter vectors a_μ, b_μ . Using eqs. 85, 87 one quickly finds

$$C = A \cdot B = c_0 \mathbb{1}_2 + i c_i \sigma_i = (a_0 b_0 - a_i b_i) \mathbb{1}_2 + i(a_0 b_i + a_i b_0 - a_k b_j \epsilon_{kji}) \sigma_i. \quad (90)$$

Matrix sum Due to the simple explicit form of exponential map, the sum of $A, B \in SU(2)$ is proportional to a further $SU(2)$ matrix:

$$C = A + B = (a_0 + b_0) \mathbb{1}_2 + i(a_i + b_i) \sigma_i = c_0 \mathbb{1}_2 + i c_i \sigma_i \quad (91)$$

with $|c|^2 = 2(a_0 b_0 + a_i b_i + 1)$ such that $\frac{C}{|c|} \in SU(2)$.

Hermitian conjugate and trace With U as in eq. 88, the Hermitian conjugate is obtained through $a_0 \rightarrow a_0, a_i \rightarrow -a_i$. The trace is simply $2a_0$ and thus notably always real. Further, eq. 89 implies

$$|a_0| = \sqrt{1 - a_i^2} \leq 1 \quad (92)$$

such that we have $-2 \leq \text{tr } U \leq 2$ for $U \in SU(2)$.

6.3 Action

To begin, we discuss the action of the general $SU(N) \otimes SU(N)$ chiral model on a D -dimensional cubic lattice with unit vectors μ and sites labelled by x . Suppose each site hosts an element ϕ of $SU(N)$ in the fundamental representation such that $\phi(x, t)$ is an $N \times N$ matrix-valued field. The defining property of such a model is that the action S is invariant under two independent global $SU(N)$ group actions. One acts on the field from the left while the other one acts from the right, making the model chiral. Specifically, we have

$$S(\phi) \rightarrow S(U\phi V^\dagger) = S(\phi), \quad U \otimes V \in SU(N) \otimes SU(N) \quad (93)$$

which, together with the usual condition that S must be real, tightly constrains the form of the action. With nearest-neighbour (NN) interactions, the $SU(N) \otimes SU(N)$ chiral model is defined by the action

$$S(\phi) = -\frac{\beta}{c} \sum_x \sum_{\mu > 0} \text{tr} (\phi^\dagger(x, t) \phi(x + \mu, t) + \phi^\dagger(x + \mu, t) \phi(x, t)) \quad (94)$$

$$= -\frac{\beta}{c} \sum_{x, \mu > 0} \text{tr} (\phi^\dagger(x, t) \phi(x + \mu, t) + \text{h.c.}) = -2\frac{\beta}{c} \sum_{x, \mu > 0} \Re \text{tr} (\phi^\dagger(x, t) \phi(x + \mu, t)) \quad (95)$$

$$= -\frac{\beta}{c} \sum_{x, y} \Re \text{tr} \{ \phi^\dagger(x, t) K_{x, y} \phi(y, t) \}. \quad (96)$$

In the final equality, the action kernel is introduced which can be identified as

$$K_{x, y} = - \sum_z \sum_{\mu > 0} (\delta_{x, z+\mu} - \delta_{x, z}) (\delta_{y, z+\mu} - \delta_{y, z}) \quad (97)$$

$$= \sum_{\mu > 0} \delta_{x, y+\mu} + \delta_{x, y-\mu} - 2\delta_{x, y}. \quad (98)$$

In the third equality adding on the Hermitian conjugate has been replaced by explicitly taking the real part of the trace. As shown in sec. 6.2, the trace of an $SU(2)$ matrix is always real, but for the sake of generality, it is left explicit in the action. The parameter β is related to the NN coupling parameter T through

$$\beta = \frac{1}{NT}. \quad (99)$$

Since the sum in x runs over all lattice sites, it is sufficient to only include the coupling along the positive direction of the lattice basis vectors. This is denoted by $\mu > 0$ and will be

left implied from here on. The constant c has its origin in the trace relation of the $SU(N)$ generators T_i . Specifically, it is shown in [35] that since $SU(N)$ is compact and simple, one can always choose a basis in which

$$\text{tr}(T_i T_j) = c \delta_{ij}. \quad (100)$$

is satisfied. Introducing c in the action causes it not to appear in the equations of motion.

Note that eq. 94 is quadratic in the field such that one might naively expect this action to describe a free field theory. However, expressing ϕ through a_μ via eq. 87 and using eq. 92 to expand a_0 in terms of a_i yields terms quartic in the scalar fields and thus a fully interacting theory. Generally, models involving a collection of scalar fields whose interactions originate entirely from the constraint to a compact space are known as non-linear sigma models.

6.4 Equations of Motion

Continuing the general discussion of the preceding section, we now derive the equations of motion for any N . Only at the stage of discretization will the discussion be narrowed to the $N = 2$ case.

6.4.1 Field

Consider the parameterisation

$$SU(N) \ni \phi(x, t) = e^{i\alpha_i(x, t)T_i} \quad (101)$$

where the parameter vector α contains the space and time dependence of the field. The momenta conjugate to the degrees of freedom $\alpha_i(x, t)$ may be defined as

$$p_i(x, t) \equiv \frac{d}{dt}\alpha_i(x, t) = \dot{\alpha}_i(x, t). \quad (102)$$

To relate these to the field $\phi(x, t)$, consider the left derivative of ϕ :

$$\frac{d}{dt}\phi(x, t) = i\dot{\alpha}_i(x, t)T_i\phi(x, t) \quad (103)$$

$$\dot{\phi}(x, t) = ip(x, t)\phi(x, t) \quad (104)$$

where in the last step the linear combination of generators $p(x, t) \equiv p_i(x, t)T_i$ was defined. Since $SU(N)$ generators are Hermitian, so will $p(x, t)$. Moving the space dependence into the subscript and suppressing the time dependence, allows to write the field equation of motion more compactly as

$$\dot{\phi}_x = ip_x\phi_x \quad \text{and} \quad \dot{\phi}_x^\dagger = -i\phi_x^\dagger p_x. \quad (105)$$

6.4.2 Momenta

In order to solve eq. 105, the equations of motion for the conjugate momenta are needed. These can be obtained by imposing conservation of energy, namely requiring

$$\frac{d}{dt}H = \dot{T} + \dot{S} = 0 \quad \text{with} \quad T = \frac{1}{2c} \sum_x \text{tr} p(x, t) p(x, t). \quad (106)$$

Inserting the definition of $p(x, t)$ reveals that the above choice for the kinetic energy is in fact such that the conjugate momenta follow independent standard Gaussians:

$$T = \frac{1}{2c} \sum_x p_i(x, t) p_j(x, t) \text{tr} T_i T_j = \frac{1}{2} \sum_x p_i(x, t) p_i(x, t). \quad (107)$$

Note that summation over i in the last expression is implied. Using the product rule and the cyclicity of the trace, we find from eq. 106 that

$$\dot{T} = \frac{1}{c} \sum_x \text{tr} p_x \dot{p}_x. \quad (108)$$

To deduce an equation for \dot{p}_x , we seek to write \dot{S} in a form similar to that of \dot{T} to identify \dot{p}_x through the conservation of energy equation. Starting from the definition of the action eq. 94 and using the field equation of motion eq. 105 as well as cyclicity yields

$$\dot{S} = -i \frac{\beta}{c} \sum_{x, \mu} \text{tr} \left\{ p_x \left(\phi_x \phi_{x+\mu}^\dagger - \phi_{x+\mu} \phi_x^\dagger \right) + p_{x+\mu} \left(\phi_{x+\mu} \phi_x^\dagger - \phi_x \phi_{x+\mu}^\dagger \right) \right\}. \quad (109)$$

As the sum runs over all lattice sites, we can relabel $x \rightarrow x + \mu$ in the second term to obtain

$$\dot{S} = -i \frac{\beta}{c} \sum_{x, \mu} \text{tr} \left\{ p_x \left(\phi_x \left(\phi_{x+\mu}^\dagger + \phi_{x-\mu}^\dagger \right) - \text{h.c.} \right) \right\} \quad (110)$$

$$= i \frac{\beta}{c} \sum_{x, \mu} \text{tr} \left\{ p_x \left((\phi_{x+\mu} + \phi_{x-\mu}) \phi_x^\dagger - \text{h.c.} \right) \right\} \quad (111)$$

which then allows to deduce the momentum equations of motion through eq. 106, 108. The result is

$$\dot{p}_x = -i\beta \sum_{\mu} (\phi_{x+\mu} + \phi_{x-\mu}) \phi_x^\dagger - \text{h.c.} . \quad (112)$$

Recalling the definition of p_x and noting that the $SU(N)$ generators are traceless, we must have $\text{tr} \dot{p}_x = 0$. This is satisfied for the $N = 2$ case as $\text{tr} \phi \in \mathfrak{R}$ but does not hold for $N \geq 3$ such that an additional term needs to be added to eq. 112 to force a vanishing trace.

At this point, the discussion will be narrowed to $N = 2$, such that $T_i = \sigma_i$ and $c = 2$. As will be shown in the following section, this also offers a significant simplification in solving the equations of motion numerically.

6.4.3 Discretisation

To use the leapfrog integrator, the momentum will be defined on a time lattice shifted by half a step size ϵ relative to the field. The discrete analogue of eq. 105 is

$$\phi(x, t + \epsilon) - \phi(x, t) = ip \left(x, t + \frac{\epsilon}{2} \right) \phi(x, t) \epsilon \quad (113)$$

$$\implies \phi(x, t + \epsilon) = \exp \left\{ ip \left(x, t + \frac{\epsilon}{2} \right) \epsilon \right\} \phi(x, t) + \mathcal{O}(\epsilon^2). \quad (114)$$

Dropping the higher-order correction reveals why this is referred to as the exponential update. Since $p(x, t) = \dot{\alpha}_i(x, t) T_i$, the above exponential is an element of $SU(N)$ for any N . Consequently, $\phi(x, t + \epsilon) \in SU(N)$, meaning the dynamics remain on the group manifold. However, for general N , the exponential must be approximated through a power series, causing ϕ to evolve off the group manifold at a rate depending on the time step ϵ . During the simulation it is therefore necessary to frequently project the fields back to the group. As shown in eq. 88, the $N = 2$ case is special since the exponential is simple to evaluate explicitly, allowing to perform the exponential update exactly such that the evolution remains within the group. A small error is always introduced due to the finite computer accuracy and requires us to check that the condition of eq. 89 remains satisfied. Doing so once every 10^4 trajectories turned out to be more than sufficient for our simulations.

Discretising the momenta equations of motion is similarly done by approximating the time derivative with a finite difference, yielding

$$p \left(x, t + \frac{\epsilon}{2} \right) = p \left(x, t - \frac{\epsilon}{2} \right) + \dot{p}(x, t) \epsilon \quad (115)$$

where one needs to employ eq. 112 on the RHS.

6.5 Fourier Acceleration

As the action kernel is similar to that of the free scalar field theory, apart from not having a mass term, the framework of sec. 4.3 applies. In particular, upon using eq. 98 and inserting the Fourier representation of the fields, eq. 96 becomes

$$S(\phi) = -\frac{\beta}{2} \frac{1}{L^2} \sum_{x, \mu} \sum_{k, k'} \Re \text{tr} \left\{ \tilde{\phi}_k^\dagger \left(2 - e^{2\pi i \frac{k' \cdot \mu}{L}} - e^{-2\pi i \frac{k' \cdot \mu}{L}} \right) \tilde{\phi}_{k'} \right\} e^{2\pi i \frac{(k' - k) \cdot x}{L}} \quad (116)$$

$$= -\frac{\beta}{2} \frac{1}{L^2} \sum_{k, k'} \sum_{\mu} \Re \text{tr} \left\{ \tilde{\phi}_k^\dagger 4 \sin^2 \pi \frac{k' \cdot \mu}{L} \tilde{\phi}_{k'} \right\} \delta_{k, k'} \quad (117)$$

$$\stackrel{!}{=} -\frac{\beta}{2} \frac{1}{L^2} \sum_{k, k'} \Re \text{tr} \left\{ \tilde{\phi}_k^\dagger \tilde{K}_{k, k'} \tilde{\phi}_{k'} \right\}. \quad (118)$$

Denoting $k \cdot \mu = k_\mu$, we thus identify the kernel in Fourier space as

$$\tilde{K}_{k, k'} = \sum_{\mu} 4 \sin^2 \pi \frac{k_\mu}{L} \quad (119)$$

which is diagonal in Fourier and $SU(2)$ space. While the former assures trivial inversion, the latter allows to commute K and its inverse with $SU(2)$ valued fields. In order for the inverse to be defined for all k , the acceleration mass parameter M is added, yielding

$$\tilde{K}_{k,k'}^{-1}(M) = \frac{1}{\sum_{\mu} 4 \sin^2 \pi \frac{k_{\mu}}{L} + M^2}. \quad (120)$$

The modified Hamiltonian is thus

$$H(\phi, p) = T(p) + S(\phi) = \frac{1}{2} \sum_{x,y} \text{tr} p_x K_{x,y}^{-1}(M) p_y + S(\phi) \quad (121)$$

$$= \frac{1}{2} \sum_{x,y} \sum_i p_i(x) K_{x,y}^{-1}(M) p_i(y) + S(\phi). \quad (122)$$

In practice, the kinetic energy is evaluated in Fourier space such that for the simulations on a two-dimensional lattice

$$H(\phi, p) = \frac{1}{2L^2} \sum_{k,k'} \sum_i \tilde{p}_i^*(k) \tilde{K}_{k,k'}^{-1}(M) \tilde{p}_i(k') + S(\phi) \quad (123)$$

$$= \frac{1}{2L^2} \frac{\sum_i |\tilde{p}_i(k)|^2}{4 \sin^2 \pi \frac{k_1}{L} + 4 \sin^2 \pi \frac{k_2}{L} + M^2} + S(\phi). \quad (124)$$

is used. Recall from eq. 46 that modifying the dynamics introduces a factor of the inverse kernel in the definition of the momenta in terms of the time derivative of the degrees of freedom. Consequently, eq. 102 becomes

$$\dot{\phi}_i(x, t) = K^{-1}(M) p_i(x, t) \quad (125)$$

and repeating the steps of eq. 104 results in the modified field equation of motion

$$\dot{\phi}_x = iK^{-1}(M) p_x \phi_x \quad \implies \quad \phi(x, t + \epsilon) = \exp \left\{ iK^{-1}(M) p \left(x, t + \frac{\epsilon}{2} \right) \epsilon \right\} \phi(x, t) \quad (126)$$

The discretisation step is analogous to that of the original equation of motion. As expected from sec. 4.3, the momentum equation of motion remains unchanged such that eq. 112 and eq. 115 remain valid. By repeating the derivation of sec. 6.4.2 and noting that the inverse kernel in eq. 125 matches the modification of the kinetic energy, this can be shown explicitly.

As the FA HMC simulations run faster than the standard HMC ones and yield smaller measurement errors, we employ the accelerated algorithm with $M = 1/10$ in the following sections if not stated otherwise. The choice of M only effects the simulation efficiency but not the physics. How to choose the acceleration mass to obtain close to optimal performance is addressed in sec. 6.9.

6.5.1 Simple Simulation Check

A simple test to conduct, but unlikely to be passed unless the simulation implementation is completely correct, is the conservation of energy. Considering the change in the Hamiltonian

ΔH between two neighbouring configurations in the chain it is expected that $\langle e^{-\Delta H} \rangle = 1$ within statistical uncertainty [13]. Simulating 2×10^3 trajectories on a $L = 16, a = 1$ lattice with $\beta = 1$ and rejecting the first 500 to assure thermalisation has occurred results in fig. 6. The computed average is indicated by the red line and the associated IAT corrected SEM is given in the legend. To within the uncertainty, the expected average is attained, giving confidence that the simulations run as expected.

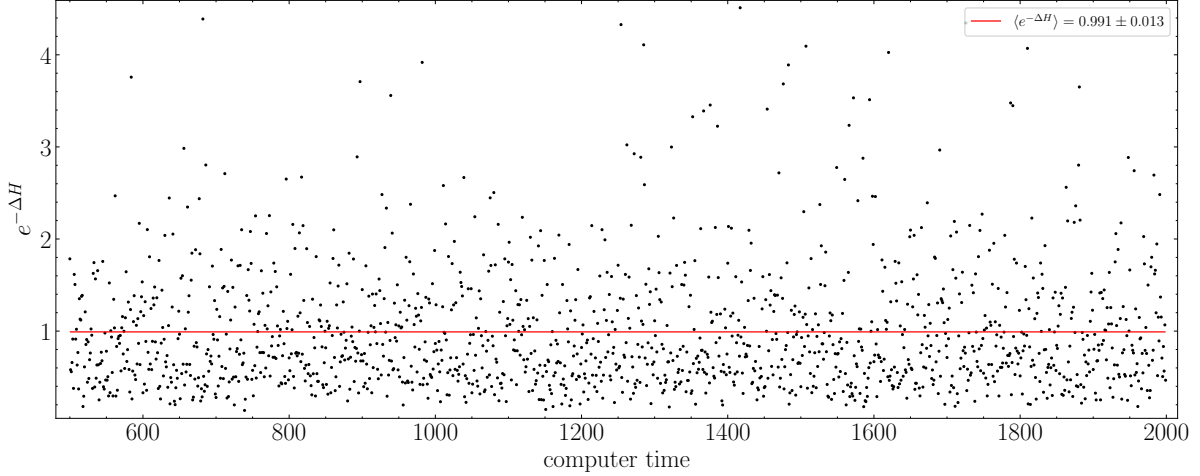


Figure 6. Exponential of the difference in the Hamiltonian between neighbouring configurations in the chain against computer time. A total of 2×10^3 trajectories are simulated using FA HMC with the first 500 declared as burn in. The lattice parameters are $L = 16, a = 1, \beta = 1$ and the computed average is indicated by the red line which, within the statistical error stated in the legend, matches the expected unit average.

6.6 Strong and Weak Coupling Expansions

Even though the considered model is not analytically solvable, expansions in powers of β and $1/\beta$ are available. Since β is defined as one over the coupling, these are referred to as the strong and weak coupling expansions respectively. A quantity whose expansions are widely discussed in the literature is the internal energy density e , allowing us to test the simulation results. For the general $SU(N) \otimes SU(N)$ model, e is commonly defined through the free energy density f

$$f = \frac{1}{VN^2} \log Z, \quad e = \frac{N}{D} \frac{\partial f}{\partial \beta} \quad (127)$$

with Z denoting the partition function of the system, while V refers to the volume of the D -dimensional lattice [5, 34]. Recalling the definition of the action from eq. 95 allows to use a standard trick of statistical physics:

$$\frac{\partial}{\partial \beta} \log Z = \frac{1}{Z} \frac{\partial}{\partial \beta} \sum_{\text{states}} \exp \left(\beta \sum_{x,\mu} \Re \text{tr} \phi_x \phi_{x+\mu}^\dagger \right) = \left\langle \sum_{x,\mu} \Re \text{tr} \phi_x \phi_{x+\mu}^\dagger \right\rangle = \left\langle -\frac{S}{\beta} \right\rangle. \quad (128)$$

Hence, for $N = 2, D = 2$ one finds

$$e = \frac{1}{VdN} \left\langle -\frac{S}{\beta} \right\rangle = \frac{1}{4V} \left\langle -\frac{S}{\beta} \right\rangle. \quad (129)$$

The normalisation of the action chosen in this report differs to that used by [5, 34], from which we draw the coupling expansion results. In order to retain the general definition of β in terms of the coupling parameter T , given in eq. 99, we absorb the relative factor into T . For the $N = 2$ case, this results in our β value being 4 times as large as its analogue in the literature.

The strong coupling expansion

$$e = \frac{1}{2}\beta + \frac{1}{6}\beta^3 + \frac{1}{6}\beta^5 + \mathcal{O}(\beta^7) \quad (130)$$

is taken from [34] while the weak coupling expansion is presented in [5]. The latter is in fact an asymptotic expansion and can be obtained by viewing mean field theory as a tree-level approximation and computing loop corrections to it [36]. Evaluating these leads to the numerical constants $Q_1 = 0.0958876$ and $Q_2 = -0.067$. Note that in [5] the internal energy density is defined as one minus eq. 129. Accounting for this yields:

$$e = 1 - \frac{3}{8\beta} \left(1 + \frac{1}{16\beta} + \left(\frac{1}{64} + \frac{3}{16}Q_1 + \frac{1}{8}Q_2 \right) \frac{1}{\beta^2} \right) + \mathcal{O}(\beta^{-4}). \quad (131)$$

The leading order behaviour is easily confirmed. Configurations are weighted by the Boltzmann factor e^{-S} such that the equilibrium will be dominated by configurations yielding the smallest action. In sec. 6.2 we showed that $-2 \leq \text{tr } \phi_x \leq 2$ for $\phi_x \in SU(2)$, meaning that the maximal contribution from the sum over nearest neighbour in eq. 95 is $2+2 = 4$. In such case the remaining sum over x produces a factor of the lattice volume, yielding $S_{min} = -4V\beta$. By eq. 129 it follows that to first order $e \approx 1$ for large β .

The simulation results as well as the coupling expansions are presented in fig. 7. For each value of β , 5×10^3 trajectories with 10% burn in are simulated on a $L = 16$ lattice with unit spacing. As the internal energy density is independent of the lattice size, choosing a small lattice is unproblematic and results in faster simulations. As FA HMC reduces autocorrelation compared to standard HMC, we obtain precise measurements even from a small number of trajectories. To better judge the agreement between the data and the expansions, the residual is plotted in fig. 17.

In their range of applicability, the predictions by the analytic expansions are matched by the simulation results at the sub-percent level. Moreover, the simulations offer results in the range $\beta \in [0.5, 1.0]$ which is not accessible by either coupling expansion. Equivalent figures already exist in the literature, such as in [34, 37]. However, they contain greater measurement errors as a local updating simulation algorithm, such as Metropolis, is used. To the best of our knowledge, this work contains the first analysis employing FA HMC.

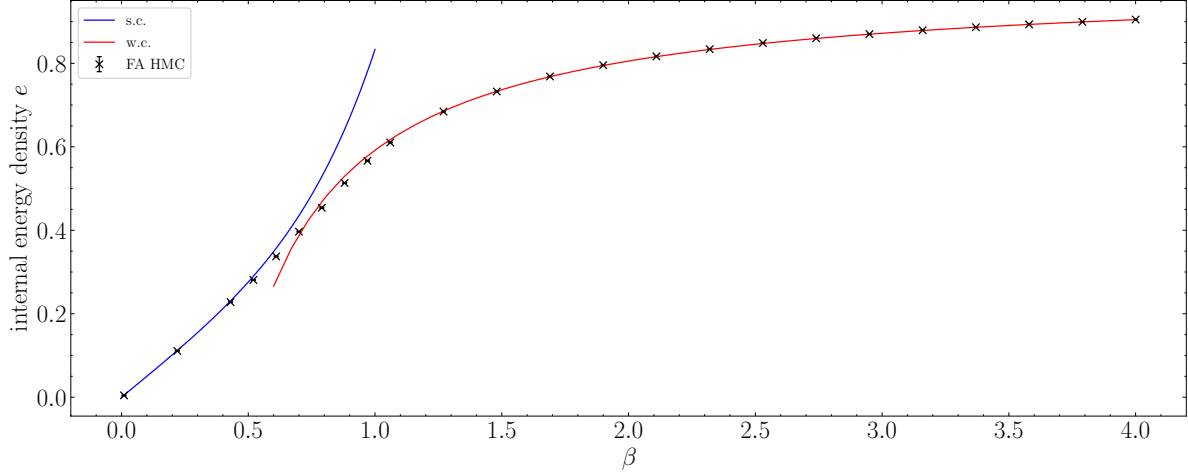


Figure 7. The internal energy density against β based on simulating 5×10^3 trajectories (10% burn in and $L = 16, a = 1$) per data point using Fourier Acceleration. The strong (blue) and weak (red) coupling expansions are superimposed and in their region of applicability agree well with the data.

6.7 Correlations

The asymptotic freedom of the $SU(2) \otimes SU(2)$ model in 2 dimensions is one of the main motivations to study this model and the numerical verification is a central objective of this report. As will be discussed in sec. 6.8.1, this requires precise measurements of the system's correlation length ξ at different values of β . The correlation length will be extracted according to the procedure described in sec. 5.1.

6.7.1 Wall to Wall correlations

However, obtaining reliable correlation function estimates to accurately infer the system's correlation length typically requires a lot of statistics. This motivates us to study wall to wall correlations, measuring the extent to which the field on two parallel planes in the lattice is correlated. As a single configuration offers many wall pairs, separated by some fixed distance d , several measurements of the correlation function can be extracted from just one configuration. Even though this reduces the number of required simulation steps, around 10^5 trajectories are required to fit ξ well. More on this later.

Consider the two blue walls with fixed separation in fig. 8, containing points with coordinates (p, i) and $(q, i + d)$ respectively. Their associated point to point correlation is defined as

$$C_{pp}(d) = \text{tr } \phi(p, i) \phi^\dagger(q, i + d) + \text{h.c.} = 2\Re \left(\text{tr } \phi(p, i) \phi^\dagger(q, i + d) \right) \quad (132)$$

where $\phi(p, i)$ is to be understood as evaluating the field at the point (p, i) on the lattice. At a fixed value of i , averaging over the row coordinates of the point pairs yields one measurement of the wall to wall correlation at separation d for the currently considered configuration. The above-mentioned advantage of wall to wall correlations is that by varying i we obtain may

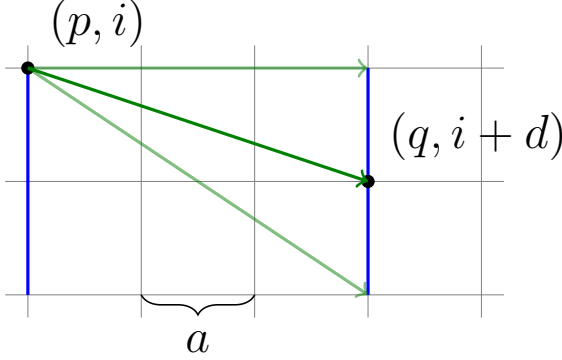


Figure 8. Composition of the wall to wall correlation function at separation $d = 3$ in terms of point to point correlations.

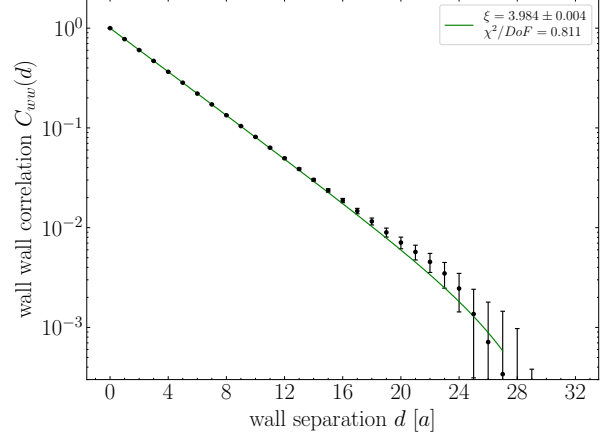


Figure 9. The wall to wall correlation function inferred from 10^5 trajectories (5% burn in) at $\beta \approx 0.8667$ on a $L = 64, a = 1$ lattice and plotted on a y logarithmic scale. A cosh function is fitted with the inferred correlation length and associated χ^2 per degree of freedom presented in the legend.

realisations of this measurement. For notational convenience, let $\phi_m(p, i)$ denote the field of the m th configuration at coordinates (p, i) . Indexing this object should be understood as accessing the components of the associated parameter vector. Denoting the ensemble average as $\langle \dots \rangle_m$ and the number of wall pairs by K , the wall to wall correlation becomes

$$C_{ww}(d) = \left\langle \frac{1}{K} \sum_i \left\{ \frac{1}{L^2} \sum_{p,q} 2\Re(\text{tr} \phi_m(p, i) \phi_m^\dagger(q, i+d)) \right\} \right\rangle_m. \quad (133)$$

Estimating the error at fixed d requires computing the autocorrelation function for each $C_{ww}(d)$ measurements chain separately.

As evident from the above description, the computation of the correlation function becomes extremely expensive for large lattice sizes, motivating to employ the FFT approach as introduced in sec. 3.2. Recall that for $A, B \in SU(2)$ and $C = AB$, we have $\text{tr} C = 2c_0 = 2(a_0b_0 - a_jb_j) \in \mathbb{R}$ with summation convention applied. Further, Hermitian conjugation leaves the 0th component of the parameter vector invariant while the spatial components are negated such that

$$C_{ww}(d) = \left\langle \frac{4}{KL^2} \sum_i \left\{ \sum_{p,q} (\phi_m(p, i))_0 (\phi_m(q, i+d))_0 + (\phi_m(p, i))_j (\phi_m(q, i+d))_j \right\} \right\rangle_m \quad (134)$$

$$= \left\langle \frac{4}{KL^2} \sum_i \left\{ \sum_{p,q} (\phi_m(p, i))_\alpha (\phi_m(q, i+d))_\alpha \right\} \right\rangle_m \quad (135)$$

$$= \left\langle \frac{4}{KL^2} \sum_i (\Phi_m(i))_\alpha (\Phi_m(i+d))_\alpha \right\rangle_m, \quad \Phi_m(i) = \sum_p \phi_m(p, i) \quad (136)$$

where summation over $\alpha = 0, 1, 2, 3$ in Euclidean space is implied. Note that for each value of α , the remaining sum is a convolution of the scalar quantity Φ_m which can be efficiently computed by the cross-convolution theorem and using Fourier transforms. Recall from sec. 3.2 that FFTs yield circular convolutions. In contrast to autocorrelations, this is desired to account for the periodic lattice boundary conditions. Further note that the constant factors may be dropped when normalising the correlation function to its value at zero separation which we will assume henceforth.

In fig. 9, the result of the above computation is shown for a lattice of size $L = 64$ with $a = 1$ and $\beta = 0.8667$, using 10^5 trajectories and rejecting the first 5% as burn in. Due to the symmetry of the correlation function, no new information is contained in the range $d \in [L/2, L]$ which has therefore been omitted in the figure. The fitted analytic prediction of the correlation function is shown in green which yields the correlation length $\xi = 3.9837(44)$. To measure the goodness-of-fit, we compute the χ_r^2 value, defined as χ^2 per degree of freedom in fit. A value close to 1 assures that the level of agreement between the data and the fitted curve is consistent with the error variance. A value much less than 1 indicates that the number of simulated trajectories is not sufficient as the error bars of neighbouring data points begin to overlap far before the noise-dominated area near the maximal separation of $L/2$ is reached. This creates the need to simulate several ten thousand trajectories to obtain an estimate of ξ accurate enough for the analysis of sec. 6.8.1.

6.7.2 Effective Mass

A further way to visualise the quality of the correlation function data is through effective mass plots. The effective mass is defined as one over the correlation length and can therefore be extracted from the analytical expected form of the correlation function⁵. Based on eq. 78 and using the identity for the sum of two cosh functions, one quickly obtains

$$m_{eff} \equiv \frac{1}{\xi} = \cosh^{-1} \left(\frac{C_{ww}(d+1) + C_{ww}(d-1)}{2 C_{ww}(d)} \right). \quad (137)$$

When plotting m_{eff} as a function of the wall separation d , one therefore expects a constant over the range where the assumed functional form of the correlation function is valid. Afterwards, noise-induced oscillations dominate. As m_{eff} depends on 3 neighbouring data points of the correlation function, noise tends to get accentuated, typically resulting in a messy plot [38].

The mass spectrum of the considered model is known and only one bound state exists [39]. Consequently, the model has a single effective mass and thus a single correlation length. For more complicated models, such as QCD, a whole tower of states exists, yielding several plateaus in an effective mass plot which correspond to different correlation lengths. As the correlation length changes with the fitting range, the fitting procedure is significantly more complicated.

⁵This generalises the process of sec. 5.1 to a cosh shaped correlation function.

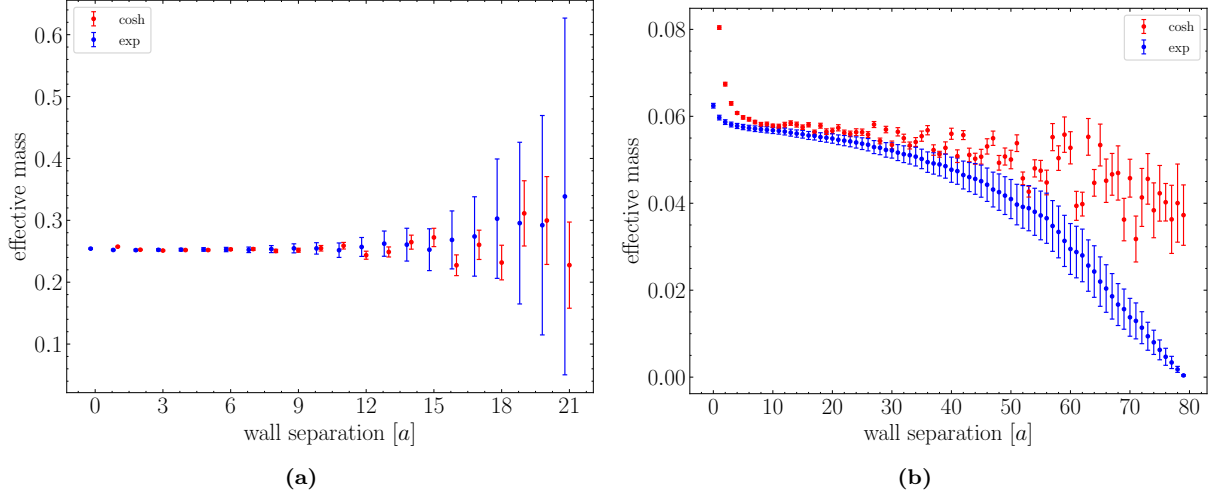


Figure 10. Effective mass plots assuming the correlation function follows a cosh (red) or a pure exponential decay (blue) for $\beta = 0.8667, L = 64$ (left) and $\beta = 1.1333, L = 160$ (right). Just under 10^5 trajectories were simulated at $a = 1$. To increase readability, the data series are slightly shifted with respect to each other. As a cosh-shaped and exponentially decaying correlation function only differ significantly for large enough arguments, the observed differences at large separations are expected. The initial drop for small separations in the right panel is a finite lattice spacing size effect and is further discussed in the main text.

The effective mass plot for simulations at $\beta = 0.8667, L = 64$ and $\beta = 1.1333, L = 160$ are shown in fig. 10. Again, $a = 1$ is used and 10^5 trajectories are simulated for either plot with a 5% burn in rejection applied. In addition to assuming a cosh as the correlation function, we also computed the effective mass for a pure exponential decay. For small separations, one expects the two data series to agree within their statistical errors before the pure exponential decay approximation breaks down. This, and the with separation increasing fluctuations, are well illustrated by fig. 10. As visible in the right panel, increasing β causes an initial decay to develop before the plateau is reached. This is observed in all effective mass plots for $\beta \gtrsim 0.9$ and is likely a finite lattice spacing effect. To confirm this, we repeated the computation for $\beta = 1.1333$ using $L = 400, a = 0.4$ which left the range of the drop measured in units of the lattice spacing unchanged. Therefore, the physical region over which the decay is observed decreased by a factor of 2.5, suggesting it vanishes in the $a \rightarrow 0$ limit. Further investigation is necessary to fully understand this phenomenon.

6.8 The β Function and Asymptotic Scaling

We are now in a position to address the physics objective of this study: asymptotic freedom. The beta function, describing the change in the coupling parameter T as the lattice spacing a varies is given by

$$\beta(T) \equiv -a \frac{\partial T}{\partial a} = -b_0 T^2 - b_1 T^3 + \mathcal{O}(T^4). \quad (138)$$

The universal coefficients b_0, b_1 are known for the general $SU(N) \otimes SU(N)$ model and presented in [5] as

$$b_0 = \frac{N}{8\pi}, \quad b_1 = \frac{N^2}{128\pi^2}. \quad (139)$$

Since the beta function is negative, it follows that increasing the parameter β in the action is equivalent to reducing the lattice spacing. Therefore, with a sufficiently large β , continuum limit results can be obtained from simulations on a finite lattice.

6.8.1 Mass over Λ Ratio

A possible approach to quantify how well the continuum limit is approximated for a given β is based on the asymptotic scaling law. During any regularisation procedure, an energy scale Λ is introduced. Asymptotic scaling requires that the dimensionless ratio of any dimensionful quantity to the appropriate power of Λ approaches a constant as the coupling strength tends to zero [5, 40]. Defining the inverse correlation length as a mass allows to test this prediction, and therefore asymptotic freedom, by computing the mass over Λ ratio. While in continuum (Quantum) Field Theories a range of equally valid regularisation prescriptions exists, numerical simulations favour lattice regularisation through the natural small distance cutoff by the lattice spacing a . The associated Λ parameter is denoted by Λ_L . In [40], Λ_L is computed at two loops (indicated by the additional subscript $2l$) and accounting for the previously mentioned factor between the definition of T yields

$$a\Lambda_{L,2l} = (4b_0T)^{-\frac{b_1}{b_0}} \exp -\frac{1}{4b_0T} \quad (140)$$

$$= (2\pi\beta)^{1/2} e^{-2\pi\beta}. \quad (141)$$

Measuring the correlation length ξ in units of the lattice spacing, the associated mass is defined as

$$M = \frac{1}{\xi a} \quad (142)$$

such that together with eq. 141 we have

$$\frac{M}{\Lambda_{L,2l}} = \frac{1}{\xi} \frac{e^{2\pi\beta}}{\sqrt{2\pi\beta}}. \quad (143)$$

Using the Bethe ansatz, it is possible to compute the ratio between the mass and Λ parameter in the modified minimal subtraction ($\overline{\text{MS}}$) scheme for the continuum model [39]. By relating the Λ parameters of the $\overline{\text{MS}}$ and L schemes, the continuum limit prediction of eq. 143 is given by

$$\frac{M}{\Lambda_{L,2l}} = \frac{32}{\sqrt{\pi e}} e^{\frac{\pi}{4}}. \quad (144)$$

Note that by combining eq. 143 and eq. 144, the β dependence of ξ reveals that the point $T = 0$ constitutes an essential singularity, such that the continuum, large β limit is an infinite order phase transition in this model.

The remaining part of this section is dedicated to testing the stability of eq. 143 with β against the above constant asymptotic scaling prediction. In order to do so, the correlation length ξ is computed for various values of β based on 10^5 trajectories (rejecting first 2×10^3 as burn in) using FA HMC with a fixed acceleration mass of $1/10$. The results are presented in tab. 1. For each value of β , the lattice size L was chosen such that $L/\xi \gtrsim 8$ to reduce systematic errors due to finite size effects as suggested by [5]. The range over which the correlation function of form eq. 78 is fitted, was determined manually. First, we inspected the correlation function plot to identify noise-dominated regions which should not be included in the fit. Adjusting the range by hand such that the value of χ_r^2 is close to 1 largely fixed the remaining freedom in the choice of fitting range. For small β , the correlation function decays into the noise quickly such that only a small number of data points are suitable for the fitting, resulting in χ_r^2 notably smaller than 1. Including the noisy regions where the fit diverges from the data artificially inflated χ_r^2 but did not result in a significant change of the inferred correlation length.

β	0.6	0.6667	0.7333	0.8	0.8667	0.9333	1.0
L	40	40	64	64	64	96	96
ξ	1.4207(17)	1.7805(19)	2.2684(20)	2.9504(29)	3.9837(44)	5.5419(60)	7.9262(96)
χ_r^2	0.204	0.268	0.387	0.243	0.811	0.766	1.019

β	1.0667	1.1333	1.2	1.2667	1.3333	1.4
L	160	160	224	400	512	700
ξ	11.407(13)	16.928(18)	25.407(30)	39.023(70)	60.94(20)	91.25(66)
χ_r^2	0.670	0.835	0.978	1.052	1.055	0.839

Table 1. The inferred correlation length with its error and the with the fit associated value of the reduced χ^2 for different pairings of β and L . The lattice size was chosen conservatively to minimise finite size effects. All simulations employ the FA HMC algorithm with the acceleration mass fixed to $1/10$ and run for 10^5 trajectories, rejecting the first 2×10^3 as burn in.

Based on this data, eq. 143 is plotted against β in fig. 11. The small uncertainty in the mass over Λ ratio is due to the precise measurement of the correlation length. For large β , we expect the ratio to approach the continuum limit prediction of eq. 144, indicated by the dashed line in fig. 11. The overall convergence to this constant is fairly poor, overshooting it by up to 14% at intermediate values of β . The behaviour near the upper bound of the considered β range suggests the onset of scaling, but firmer conclusions require additional data points at even larger values of β . It should be noted that the simulation time increases with β and with our current implementation, collecting the $\beta = 1.4$ data takes 16 days. Even though this analysis has not been performed for the $N = 2$ model in 2 dimensions, results for the $N = 3, 6, 9, 15$ cases are given in [5]. The behaviour of the mass over Λ ratio in these models is very similar to that described above. Notably, overshooting of around 20% occurs across a similar region in β and is believed to be caused by a dip in the lattice beta function. A possible approach to widen the range of asymptotic scaling is to employ improved coupling techniques in which the coupling parameter is redefined. In [5], by defining T in terms of the internal energy density, the rate of convergence was greatly improved for

all considered models. We applied the same scheme to our model but did not observe any improvements as illustrated in fig. 18. It is not clear to us what properties of the $N = 2$ model cause this technique to fail. Further investigations in that direction as well as other approaches for increasing the range of asymptotic scaling are left as future work.

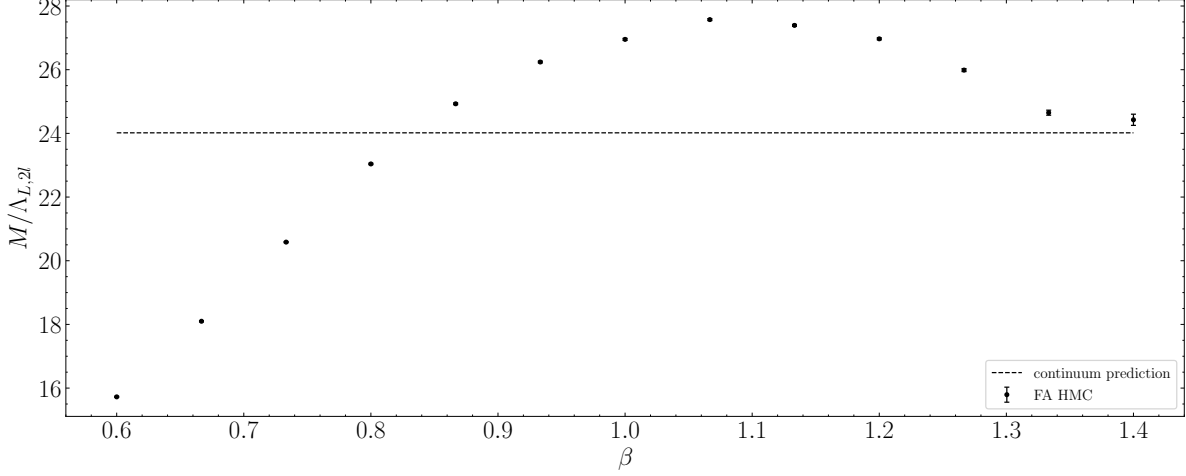


Figure 11. The mass over Λ ratio of eq. 143 against β . Each data point is the result of fitting the correlation length to the averaged correlation function data based on 10^5 trajectories (rejecting the first 2×10^3 as burn in). All simulations are performed with $a = 1$ while the lattice size L is specified in tab. 1. The dashed line indicates the continuum limit prediction of eq. 144.

6.9 Critical Slowing Down

Finally, the algorithmic objective of the project is addressed by quantifying the simulation efficiency improvement by employing Fourier Acceleration. Before constructing the ratio of the standard HMC and FA HMC cost function, defined in eq. 45, an observable to measure the IAT of needs to be chosen. The susceptibility χ is particularly well suited for this purpose as it is dominated by slow modes, resulting in a quickly growing autocorrelation time near criticality [6]. For a lattice of volume V , the susceptibility of a single configuration is defined as

$$\chi = \frac{1}{V} \sum_{x,y} \Re \text{tr} \phi_x \phi_y = \frac{1}{2V} \sum_{x,y} \text{tr} \phi_x \phi_y + \text{h.c.} \quad (145)$$

and can be viewed as the average point-to-point correlation of that configuration. Applying the ideas introduced in sec. 6.7.1 allows to significantly improve the speed at which χ is computed. Analogous to eq. 136, with $\alpha = 0, 1, 2, 3$ and $\Phi(i) = \sum_p \phi(p, i)$, one finds

$$\chi = \frac{2}{L^2} \sum_j \sum_i \left\{ \sum_{p,q} \phi(p, i)_\alpha \phi(q, j)_\alpha \right\} = \frac{2}{L^2} \sum_j \left\{ \sum_i (\Phi(i))_\alpha (\Phi(j))_\alpha \right\}. \quad (146)$$

Due to the lattice periodicity and since i and j both index columns of the lattice, we can always express i as a shifted version of j . Therefore, the inner bracket is just the sum

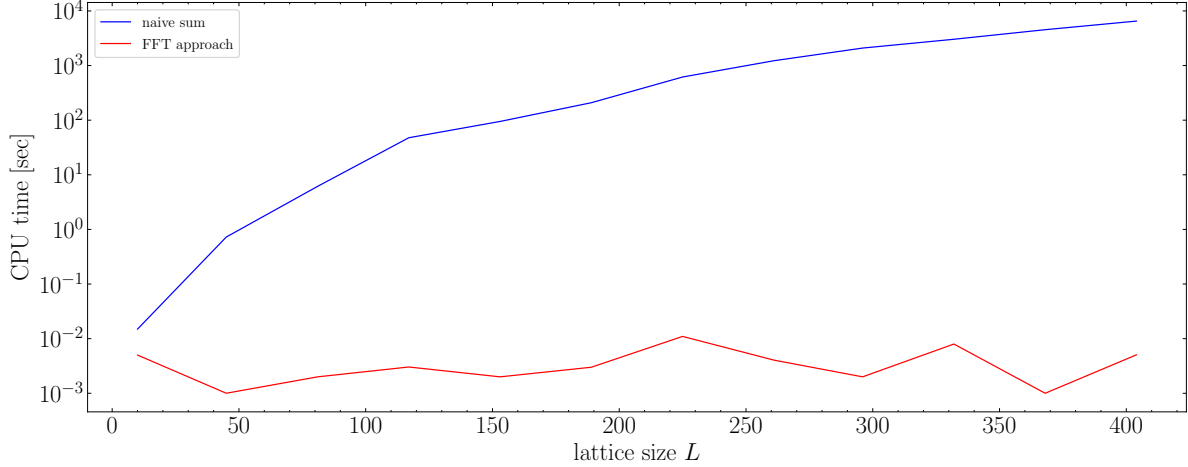


Figure 12. Comparing the CPU time needed to compute the susceptibility of a single configuration using eq. 145 (blue) and eq. 146 (red) on a logarithmic y scale. The total number of samples to correlate is given by $4L^2$ such that in extreme cases the FFT approach runs over 6 orders of magnitudes faster. The limited clock precision masks the scaling of the FFT method.

of 4 scalar convolutions and has the same form as the correlation function for a single configuration (c.f. eq. 136). The remaining sum over j is thus simply performed by summing up the resulting correlation function.

In fig. 12 the CPU time required to compute χ based on performing the double sum in eq. 145 explicitly and through eq. 146 using FFTs are compared for varied sizes of the two-dimensional lattice. The latter approach offers a run time reduction of almost 6 orders of magnitude near the upper limit of the survey. Considering that to test Fourier Acceleration the susceptibility will be computed for 10^5 configurations per data point at even larger values of L , this speed-up is vital to attain a reasonable execution time. Note that the limited precision of the machine clock and background processes are likely to dominate over the L dependence in the FFT approach.

Calibrating Acceleration Mass In order to realise the full potential of FA HMC, the acceleration mass parameter M must be tuned. While the optimal value of M is not obvious for interacting theories, a common choice is the effective mass [15]. To validate this, we measured the cost function of eq. 45 for various values M at $\beta = 1.1333$ with $L = 160, a = 1$ based on 1.5×10^4 trajectories with 10% burn in. The findings are presented in fig. 13 where the cost function is normalised to its value at $M = 1/\xi$.

Even though simulation performance increased by around 10% when using $M = 0.02$ instead of $M = 1/\xi$, conducting such a grid search for each value of β is computationally expensive. We, therefore, adapt the choice of $M = 1/\xi$ for the following analysis. We also conducted a second search including larger values of M for which the cost function was observed to increase monotonically. This is expected since a large value of M tends to dominate any \mathbf{k} dependence in the inverse kernel which causes acceleration in the first place.

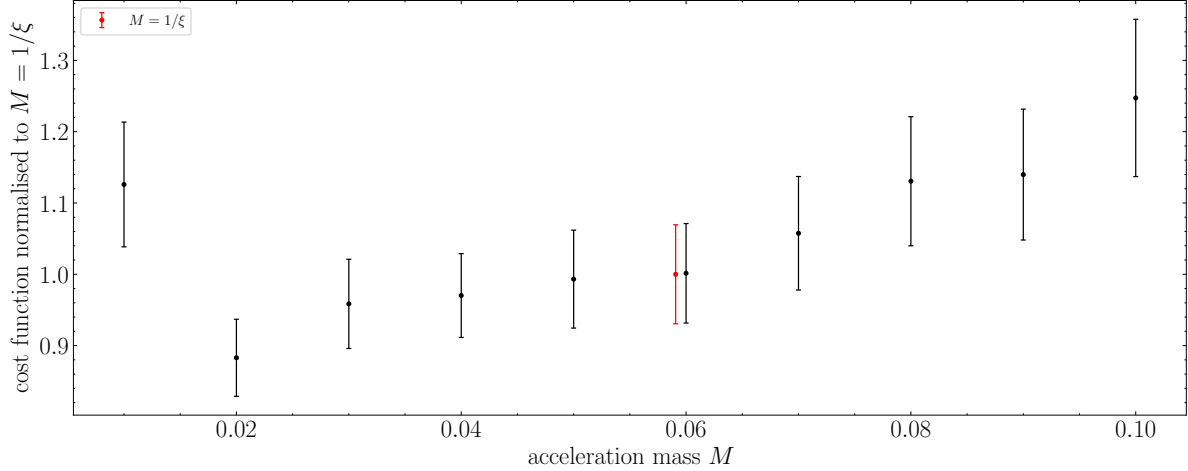


Figure 13. The cost function of eq. 45 for different values of the acceleration mass, normalised to the conventional choice $M = 1/\xi$. Each data point corresponds to 1.5×10^4 trajectories (with 10% burn in) for $L = 160, a = 1, \beta = 1.1333$. The overall shape of the cost function is shallow in the considered region such that $M = 1/\xi$ yields acceptable, but not quite optimal acceleration.

For each value pair β, L from tab. 1 we simulated 10^5 trajectories using both standard HMC and FA HMC to deduce the susceptibility IAT in either case. The first 5×10^3 trajectories are discarded as burn in. The data, as well as the fitted power law of eq. 39, are plotted in fig. 14. First consider standard HMC, where the dynamical critical exponent was inferred to be 1.777 ± 0.037 . As expected, the fluctuations in the IAT increase notably for HMC towards the upper end of the survey. By doubling the amount of burn in, we excluded the possibility that these are due to the chains not having fully equilibrated. In order to improve the quality of these data points, the statistics must be increased. Due to long simulation run times, this was not possible in the given time constraints. It is unexpected to find that for the smallest correlation length, HMC appears to yield a smaller IAT compared to FA HMC. As we are mainly interested in the large ξ behaviour, we did not investigate this phenomenon any further. For FA HMC, we found a critical exponent of 0.21 ± 0.01 . That critical slowing down was not fully avoided is likely due to the non-optimal choices of the acceleration masses. Nevertheless, FA HMC still offers significant improvements compared to standard HMC. In particular, employing Fourier Acceleration allows to gain a factor of roughly 320 in the susceptibility autocorrelation time at $\beta = 1.4, \xi \approx 91.25$. Consequently, one can expect that HMC requires 300 times more trajectories than FA HMC to achieve comparable accuracy.

Finally, we compute the HMC and FA HMC cost function ratio which is shown in fig. 15. Similar to the QHO case (c.f. fig. 5), the ratio follows a power law whose exponent was inferred to be 0.744 in the present case. This is roughly equal to half the difference between the previously found dynamical critical exponents, suggesting the cost function behaviour is dominated by the square root of the IAT. Manual inspection of the cost function factors confirmed this. Again, the unexpected result that HMC slightly outperforms FA HMC for small ξ was not further investigated. The main conclusion is that FA HMC proved to be up to 20 times more cost-efficient than HMC for the $SU(2) \otimes SU(2)$ model.

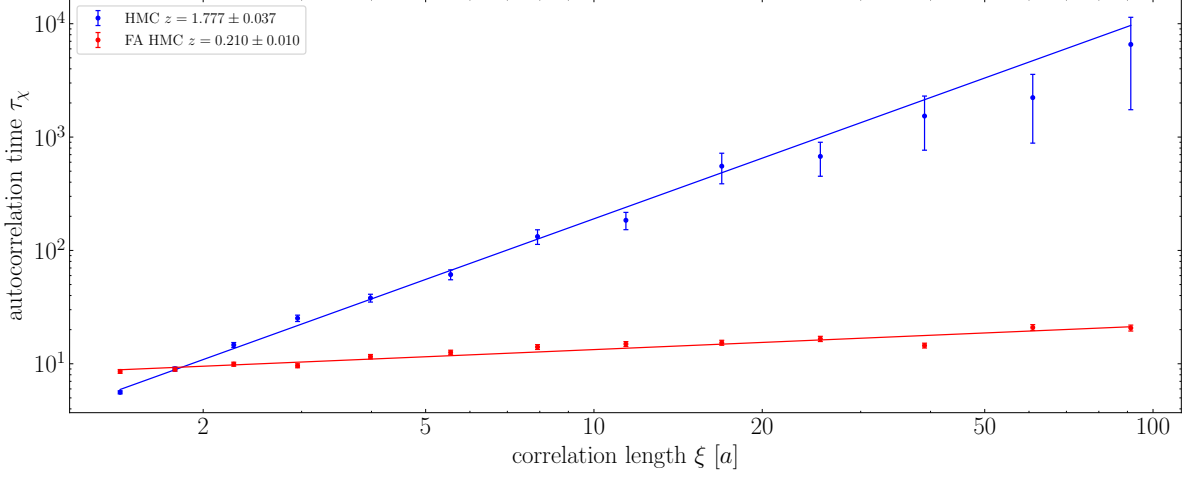


Figure 14. The integrated autocorrelation time of the susceptibility for the $SU(2) \otimes SU(2)$ model against the system correlation length ξ on log-log axes for standard HMC (blue) and FA HMC (red) simulations. Fitted power laws are superimposed with the value of the critical exponent presented in the legend. Each data point corresponds to 10^5 trajectories with the lattice size L and the value of β as specified in tab. 1.

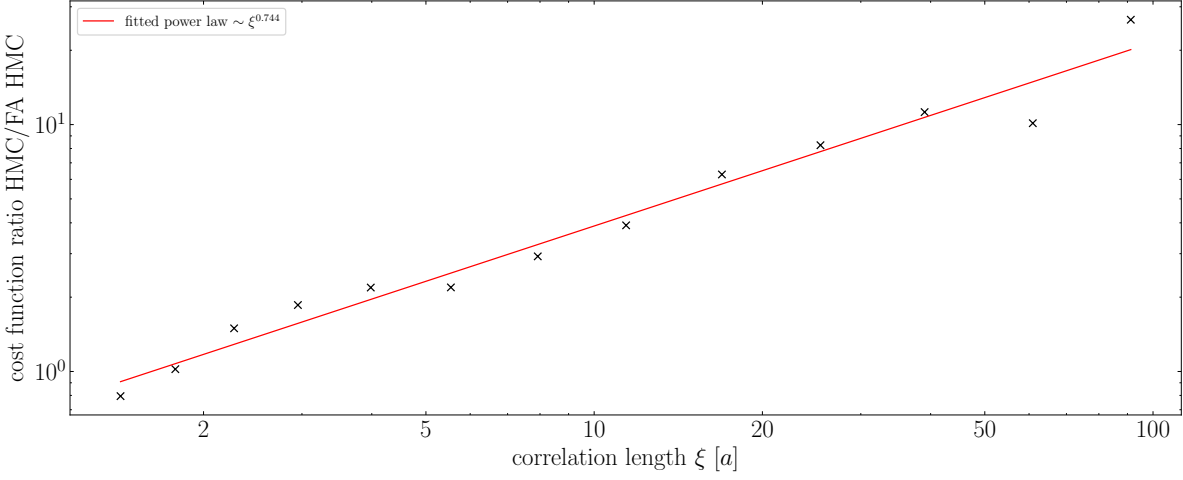


Figure 15. Ratio of the standard HMC and FA HMC cost functions against the system correlation length on a log-log scale with a fitted power law superimposed. The simulation parameters are identical to those used in fig. 14.

7 Conclusion

The increased available computing power over the last few decades allowed for ever more precise lattice simulations. To this date, the lattice approach appears to be the most promising path to better understand confinement and in general the non-perturbative nature of QCD at low energies. Considering the $SU(N) \otimes SU(N)$ principal chiral models presents a testbed for the study of non-abelian gauge theories, as they are computationally easier to manage while offering sufficiently similar physics, in the sense of being fully interacting and asymptotically free quantum field theories. The hope is to utilise the developed techniques and gained insights from studying the simpler models to improve our understanding of gauge theories.

We began by providing an introduction to Monte Carlo techniques and error analysis. This is followed by a general discussion of Fourier Acceleration before using the one-dimensional quantum harmonic oscillator as a paradigm to illustrate this technique.

The main focus of the project however lay in the numerical simulation of the two-dimensional $SU(2) \otimes SU(2)$ model. We computed the internal energy density depending on the value of the coupling and found the numerical results to agree with analytic expansions at a level of precision exceeding that of existing studies. The novel employment of the Fourier accelerated Hamiltonian Monte Carlo algorithm for the analysis made this possible. In addition, we numerically investigated asymptotic freedom which has not been done previously for this theory. The findings agree with other $SU(N) \otimes SU(N)$ studies. However, further investigation into the negative result from the improved coupling prescription is required to refine the deduction that only the onset of asymptotic freedom is observed.

The algorithmic objective of quantifying the degree to which critical slowing down can be avoided by integrating Fourier Acceleration into our Hamiltonian Monte Carlo simulations has been fully achieved by measuring an improvement of a factor of 320. For the considered model, such analysis is original to the literature. We identified a variety of avenues for future work, including optimising the trajectory length and improving the acceleration mass choice. Both offer the opportunity to further increase the simulation performance, making computations closer to the continuum more accessible such that more precise measurements can be made at fixed CPU time.

To achieve these objectives, we developed a simulation and analysis package which can serve as a foundation for further studies of $SU(N) \otimes SU(N)$ principal chiral models. It will be interesting to see how the performance of Fourier Acceleration compared to standard HMC develops against N . In addition, we developed an efficient approach to compute (auto-)correlation functions, resulting in a negligible time requirement for data analysis. It will be intriguing to see the effect of this method within and beyond the domain of lattice computations.

References

- [1] Antonin Portelli. *Lattice Field Theory*. lecture notes, Gauge Theories in Particle Physics PGPH11099. The University of Edinburgh, Mar. 2023.
- [2] Babak Abi et al. *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*. In: *Physical Review Letters* 126.14 (Apr. 2021). arXiv: 2104.03281.
- [3] Tatsumi Aoyama et al. In: 887 (Dec. 2020), pp. 1–166. arXiv: 2006.04822.
- [4] Constantia Alexandrou et al. *Lattice calculation of the short and intermediate time-distance hadronic vacuum polarization contributions to the muon magnetic moment using twisted-mass fermions*. 2022. arXiv: 2206.15084.
- [5] Paolo Rossi and Ettore Vicari. *Two-dimensional $SU(N) \times SU(N)$ chiral models on the lattice*. In: *Phys. Rev. D* 49.3 (Feb. 1994). [Erratum: *Phys.Rev.D* 55, 1698 (1997)], pp. 1621–1628. arXiv: hep-lat/9307014.
- [6] Elbio R.A. Dagotto and John B. Kogut. *Testing accelerated stochastic algorithms in two dimensions: The $SU(3) \times SU(3)$ spin model*. In: *Nuclear Physics B* 290 (Jan. 1987), pp. 451–468. ISSN: 0550-3213.
- [7] Martin Hasenbusch and Steffen Meyer. *Multigrid acceleration for asymptotically free theories*. In: *Phys. Rev. Lett.* 68.4 (Jan. 1992), pp. 435–438.
- [8] Julian Wack. *MPhysProject*. Version v1.0.0. Mar. 2023.
- [9] Michael J. Creutz and Barry A. Freedman. *A statistical approach to quantum mechanics*. In: *Annals of Physics* 132.2 (Apr. 1981), pp. 427–462. ISSN: 0003-4916.
- [10] Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng. *Handbook of Markov Chain Monte Carlo*. Chapman and Hall/CRC, 2011. ISBN: 1420079417.
- [11] Christof Gattringer and Christian B. Lang. *Quantum chromodynamics on the lattice: An introductory presentation*. Springer, 2010. ISBN: 9783642018497.
- [12] Michael Betancourt. *A Conceptual Introduction to Hamiltonian Monte Carlo*. 2017. arXiv: 1701.02434.
- [13] Anosh Joseph. *Markov Chain Monte Carlo Methods in Quantum Field Theories*. Springer International Publishing, 2020. ISBN: 978-3-030-46044-0. arXiv: 1912.10997.
- [14] Nicholas Metropolis et al. *Equation of State Calculations by Fast Computing Machines*. In: *The Journal of Chemical Physics* 21.6 (June 1953), pp. 1087–1092. ISSN: 0021-9606.
- [15] Simon Duane, Richard Kenway, Brian J. Pendleton, and Duncan Roweth. *Acceleration of gauge field dynamics*. In: *Physics Letters B* 176.1 (Aug. 1986), pp. 143–148. ISSN: 0370-2693.
- [16] Samuel Livingstone, Michael Betancourt, Simon Byrne, and Mark Girolami. *On the Geometric Ergodicity of Hamiltonian Monte Carlo*. 2016. arXiv: 1601.08057.
- [17] Herbert Goldstein, Charles P. Poole, and John L. Safko. *Classical mechanics*. 3rd ed. Addison Wesley, 2002. ISBN: 0201657023.

- [18] Jack Frankland. *Accelerated Tempering Dynamics in HMC Simulations of Lattice Field Theory*. MPhys project report. Visited on 16/03/2023. The University of Edinburgh, Apr. 2018.
- [19] Benedict Leimkuhler and Sebastian Reich. *Simulating Hamiltonian Dynamics*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2005. ISBN: 9780521772907.
- [20] Alan V. Oppenheim, Ronald W. Schafer, and John R. Buck. *Discrete-Time Signal Processing*. Second. Prentice-Hall, 1999. ISBN: 0137549202.
- [21] Sergio Caracciolo and Alan D. Sokal. *Dynamic critical exponent of some Monte Carlo algorithms for the self-avoiding walk*. In: *Journal of Physics A: Mathematical and General* 19.13 (Sept. 1986), p. L797. ISSN: 0305-4470.
- [22] Neal Madras and Alan D. Sokal. *The pivot algorithm: A highly efficient Monte Carlo method for the self-avoiding walk*. In: *Journal of Statistical Physics* 50.1 (Jan. 1988), pp. 109–186. ISSN: 1572-9613.
- [23] Ulli Wolff. *Monte Carlo errors with less errors*. In: *Comput. Phys. Commun.* 156 (2004). [Erratum: *Comput.Phys.Comm.* 176, 383 (2007)], pp. 143–153. arXiv: hep-lat/0306017.
- [24] Eric W. Weisstein. *Cross-Correlation Theorem*. From *MathWorld*—A Wolfram Web Resource. Visited on 14/11/2022.
- [25] William Press, Saul Teukolsky, William Vetterling, and Brian Flannery. *Numerical Recipes in Fortran: The Art of Scientific Computing*. 2nd ed. Cambridge University Press, 1992. ISBN: 9780521430647.
- [26] Guido Cossu et al. *Testing algorithms for critical slowing down*. In: *EPJ Web of Conferences* 175.02008 (2018). arXiv: 1710.07036.
- [27] James Doherty. *The sine-Gordon Model with Fourier Acceleration*. MPhys project report. The University of Edinburgh, Mar. 2021.
- [28] Simon Duane and Brian J. Pendleton. *Gauge invariant Fourier acceleration*. In: *Physics Letters B* 206.1 (May 1988), pp. 101–106. ISSN: 0370-2693.
- [29] Tuan Nguyen et al. *Riemannian manifold hybrid Monte Carlo in lattice QCD*. 2021. arXiv: 2112.04556.
- [30] Aleksandra Slapik and Willian Serenone. *Lattice Monte Carlo Study of the Harmonic Oscillator in the Path Integral Formulation*. Summer Student Report. DESY, 2012.
- [31] Ronnie Rodgers and Laura Raes. *Monte Carlo simulations of harmonic and anharmonic oscillators in discrete Euclidean time*. Summer Student Report. DESY, 2014.
- [32] Alexander A. Migdal. *Recursion Equations in Gauge Field-Theories*. In: *30 Years of the Landau Institute - Selected Papers*. World Scientific Publishing, 1996, pp. 114–119.
- [33] Frederic Green and Stuart Samuel. *Chiral models: Their implication for gauge theories and large N* . In: *Nuclear Physics B* 190.1 (July 1981), pp. 113–150. ISSN: 0550-3213.

- [34] Arunabha Guha and S.-C. Lee. *Improved mean field studies of $SU(N)$ chiral models and comparison with numerical simulations*. In: *Nuclear Physics B* 240.2 (Sept. 1984), pp. 141–170. ISSN: 0550-3213.
- [35] Neil Turok. *Symmetries of Particles and Fields*. lecture notes, Symmetries of Particles and Fields PGPH11097. The University of Edinburgh, Sept. 2022.
- [36] Yves J. Brihaye and Paolo Rossi. *The weak-coupling phase of lattice spin and gauge models*. In: *Nuclear Physics B* 235.2 (June 1984), pp. 226–258. ISSN: 0550-3213.
- [37] Arunabha Guha and S.-C. Lee. *Stochastic quantization for numerical simulation*. In: *Physics Letters B* 134.3 (Jan. 1984), pp. 216–220. ISSN: 0370-2693.
- [38] George T. Fleming, Saul D. Cohen, Huey-Wen Lin, and Victor Pereyra. *Excited-state effective masses in lattice QCD*. In: *Phys. Rev. D* 80.7 (Oct. 2009), pp. 1–13. arXiv: 0903.2314.
- [39] Janos Balog, Saurabh Naik, Ferenc Niedermayer, and Peter H Weisz. *Exact mass gap of the chiral $SU(n) \times SU(n)$ model*. In: *Phys. Rev. Lett.* 69.6 (Aug. 1992), pp. 873–876.
- [40] Junko Shigemitsu and John B. Kogut. *A study of Λ parameters and crossover phenomena in $SU(N) \times SU(N)$ sigma models in two dimensions*. In: *Nuclear Physics B* 190.2 (Aug. 1981), pp. 365–411. ISSN: 0550-3213.
- [41] James Binney and David Skinner. *The Physics of Quantum Mechanics*. Oxford University Press, 2013. ISBN: 9780199688579.

Appendices

Note — Appendices are provided for completeness only and any content included in them will be disregarded for the purposes of assessment.

A Additional Proofs

A.1 Liouville's Theorem

One approach to proving this property is to show that the Jacobian J of the infinitesimal mapping $\mathcal{H}_\delta : t \rightarrow t + \delta t$ is $1 + \mathcal{O}(\delta t^2)$. Specifically, Taylor expanding to first order yields

$$J = \begin{vmatrix} \frac{\partial}{\partial \phi_j} \phi_i(t + \delta t) & \frac{\partial}{\partial \pi_j} \phi_i(t + \delta t) \\ \frac{\partial}{\partial \phi_j} \pi_i(t + \delta t) & \frac{\partial}{\partial \pi_j} \pi_i(t + \delta t) \end{vmatrix} = \begin{vmatrix} \delta_{ij} + \frac{\partial \phi_i}{\partial \phi_j} \delta t & \frac{\partial \phi_i}{\partial \pi_j} \delta t \\ \frac{\partial \pi_i}{\partial \phi_j} \delta t & \delta_{ij} + \frac{\partial \pi_i}{\partial \pi_j} \delta t \end{vmatrix} + \mathcal{O}(\delta t^2) \quad (147)$$

$$= \det(\mathbb{1} + A\delta t) + \mathcal{O}(\delta t^2) = 1 + \delta t \operatorname{tr} A + \mathcal{O}(\delta t^2) = 1 + \mathcal{O}(\delta t^2). \quad (148)$$

where Hamilton's equations assure that the trace of the matrix A vanishes. Consequently, phase space volumes are constants in time:

$$V(t + \delta t) = \int dV(t + \delta t) = \int dV(t) J = V(t) + \mathcal{O}(\delta t^2) \quad (149)$$

$$\implies \frac{dV}{dt} = \lim_{\delta t \rightarrow 0} \frac{V(t + \delta t) - V(t)}{\delta t} = \lim_{\delta t \rightarrow 0} \mathcal{O}(\delta t) = 0. \quad (150)$$

A.2 Cross Correlation Theorem for Discrete Variables

We define the inverse Fourier transform of a discrete variable $x_j, j = 0, 1 \dots L - 1$ as

$$x_j = \mathcal{F}^{-1}[\tilde{x}] = \frac{1}{L} \sum_{k=0}^{L-1} \tilde{x}_k e^{2\pi i \frac{kj}{L}}. \quad (151)$$

The real space convolution of two such variables x, y is then given by the product in Fourier space as

$$\mathcal{F}^{-1}[(\mathcal{F}[x])^* \mathcal{F}[y]]_j = \frac{1}{L} \sum_k \mathcal{F}[x]_k^* \mathcal{F}[y]_k e^{2\pi i \frac{kj}{L}} = \frac{1}{L} \sum_k \sum_m \sum_n x_m^* y_n e^{2\pi i \frac{k(j+m-n)}{L}} \quad (152)$$

$$= \sum_m \sum_n x_m^* y_n \delta_{n,j+m} = \sum_m x_m^* y_{m+j} \quad (153)$$

$$= x * y. \quad (154)$$

B Quantum Harmonic Oscillator

In this section, we give some of the analytic results quoted in the numerical study of the one-dimensional quantum harmonic oscillator. In addition, the results for the ground state energy and wave function will be presented.

B.1 Analytic Results

The following results are derived in [9, 18] for the special case of a unit mass particle in a potential $V(x) = \frac{1}{2}\mu^2 x^2$. Using $\mu = \sqrt{m}\omega$ allows to map the results to the potential considered in this report. For notational convenience, define

$$A = \omega \sqrt{1 + \frac{1}{4}a^2\omega^2}, \quad R = \sqrt{1 + a^2A^2} - aA. \quad (155)$$

The lattice theory predictions and their continuum limit, i.e. the $L \rightarrow \infty$ and $a \rightarrow 0$ limits, are

$$\langle x^2 \rangle = \frac{1}{2mA} \frac{1 + R^N}{1 - R^N} \rightarrow \frac{1}{2m\omega} \quad (156)$$

$$E_0 = \frac{2m}{a^2} \left(\sqrt{1 + a^2A^2} - 1 \right) \langle x^2 \rangle \rightarrow \frac{1}{2}\omega \quad (157)$$

$$\psi_0(x) = \left(\frac{mA}{\pi} \right)^{1/4} \exp \left(-\frac{1}{2}mA x^2 \right) \rightarrow \left(\frac{m\omega}{\pi} \right)^{1/4} \exp \left(-\frac{1}{2}m\omega x^2 \right) \quad (158)$$

$$E_1 - E_0 = -\frac{1}{a} \ln \frac{R^2 - R^{N-2}}{R - R^{N-1}} \rightarrow \omega. \quad (159)$$

Note that in sec. 5.1 only one of these limits is considered resulting in finite size effects. Specifically, L was fixed while taking $a \rightarrow 0$, causing the numerical and analytic lattice predictions to diverge as seen in fig. 3 and fig. 16.

B.2 Numerical Ground State Results

Once the Markov chain of position measurements has been obtained, the squared modulus of the ground state wave function $|\psi_0(x)|$ (i.e. the particle's spatial probability density) is obtained by computing a normalised histogram for the collected position data. In order to obtain an error estimate, we split the shuffled data into 100 equally sized subsets and compute the histogram using 100 bins. The collection of bin heights then allows to infer the average as well as the standard error of the mean. The result from simulating 10^4 trajectories (with 10% burn in rejected) at $L = 100, m = 1, \omega = 1$, along with the predictions of eq. 158 are shown in the left panel of fig. 16. Overall, the data is in good agreement with the lattice prediction even though only a relatively small number of trajectories was simulated, underlining the efficiency of the HMC algorithm. The deviations to the continuum are due to the relatively

large value of the considered lattice spacing. Note that the binning procedure introduces a systematic error which is likely to be the cause of the irregular behaviour around $x = 0$. The choice for the number of splits is a further source of systematic errors and the above approach is motivated to yield on average L measurements per bin.

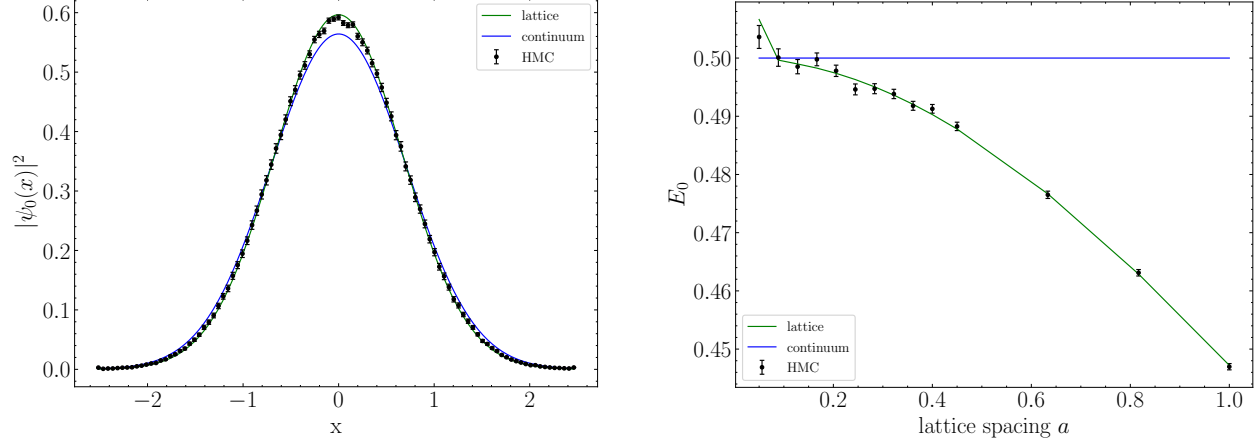


Figure 16. Comparison of the numerical results for the quantum harmonic oscillator ground state with the analytic predictions. In both cases $L = 100, m = 1, \omega = 1$ is used. Left: The wave function inferred from 10^4 trajectories with the binning procedure laid out in the main text. Right: The ground state energy deduced through the virial theorem against the lattice spacing from 5×10^5 trajectories.

Using the virial theorem of classical dynamics, which can be shown to hold in quantum mechanics [41] as well, allows estimating the ground state energy. Specifically, for one-dimensional dynamics with kinetic energy T and potential $V(x)$ with derivative $V'(x)$ one finds

$$2\langle T \rangle = \langle xV'(x) \rangle. \quad (160)$$

For $V(x) = 1/2m\omega^2x^2$, this yields

$$E_0 = \langle 0|T + V|0 \rangle = \left\langle \frac{1}{2}xV'(x) + V(x) \right\rangle = m\omega^2 \langle x^2 \rangle. \quad (161)$$

The numerical results and the analytic predictions of eq. 157 based on 10^4 trajectories (with the first 10% rejected as burn in) and $L = 100, m = 1, \omega = 1$ are shown in the right panel of fig. 16. Further, the increasing error size with decreasing a is a result of the increasing autocorrelations and is therefore also expected. Employing Fourier Acceleration or simply simulating more trajectories allows us to reduce these errors. Nevertheless, the numerical results agree with the theoretical predictions to a satisfactory level and show the desired convergence to the continuum result before diverging due to finite size effects.

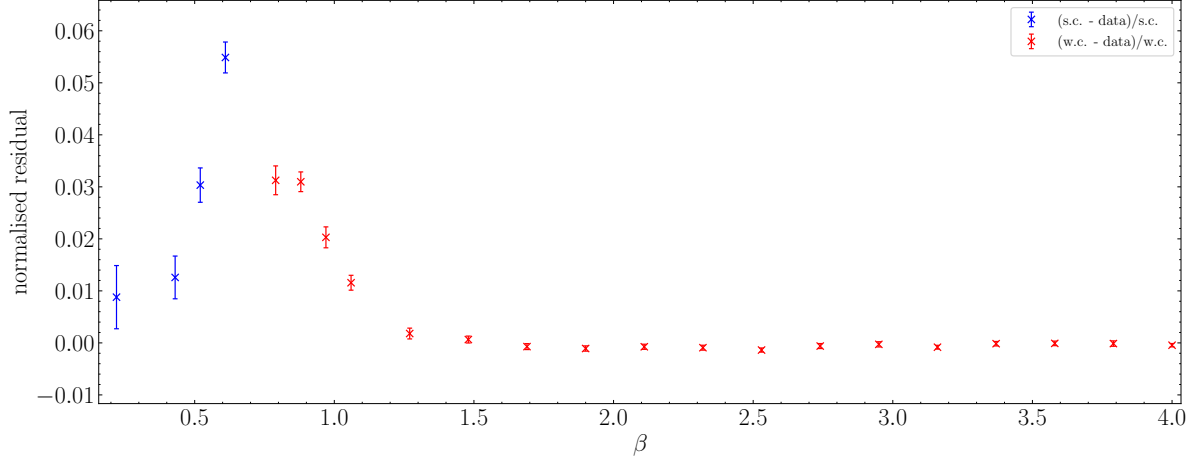


Figure 17. The normalised residual between the FA HMC results and the strong and weak coupling expansions. 5×10^3 trajectories are simulated per data point with a 10% burn in rejection applied. The first point of the strong coupling expansion is noise-dominated and thus omitted. The simulations are run on a $L = 16, a = 1$ lattice.

C $SU(2) \otimes SU(2)$

C.1 Coupling Expansion Residuals

In fig. 17 the normalised residuals between the strong and weak coupling expansions and the simulation results are shown. Note that the first data point of the strong coupling expansion is not plotted as both the analytic prediction and the numerical result are close to zero such that the residual is dominated by noise and becomes large. In the range $\beta \in [0.5, 1]$ the simulations offer corrections of up to 6%. For large values of β , the high level of agreement to the weak coupling expansion becomes apparent.

C.2 Improved Coupling

Considering a general $SU(N) \otimes SU(N)$ model and following [5], we define a new coupling parameter to be proportional to the internal energy density e (defined in eq. 129):

$$T_e = \frac{8N}{N^2 - 1}e, \quad \beta_e = \frac{1}{NT_e}. \quad (162)$$

The relation of the Λ parameter of this new scheme to the lattice scale Λ_L is known to be

$$\frac{\Lambda_e}{\Lambda_L} = e^{\pi \frac{N^2 - 2}{4N^2}} \quad (163)$$

while, due to the universality of the first two beta function coefficients, the two-loop function is given by

$$\Lambda_{L,2l}(4x) = \Lambda_{e,2l}(x) = \sqrt{8\pi x} e^{-8\pi x}. \quad (164)$$

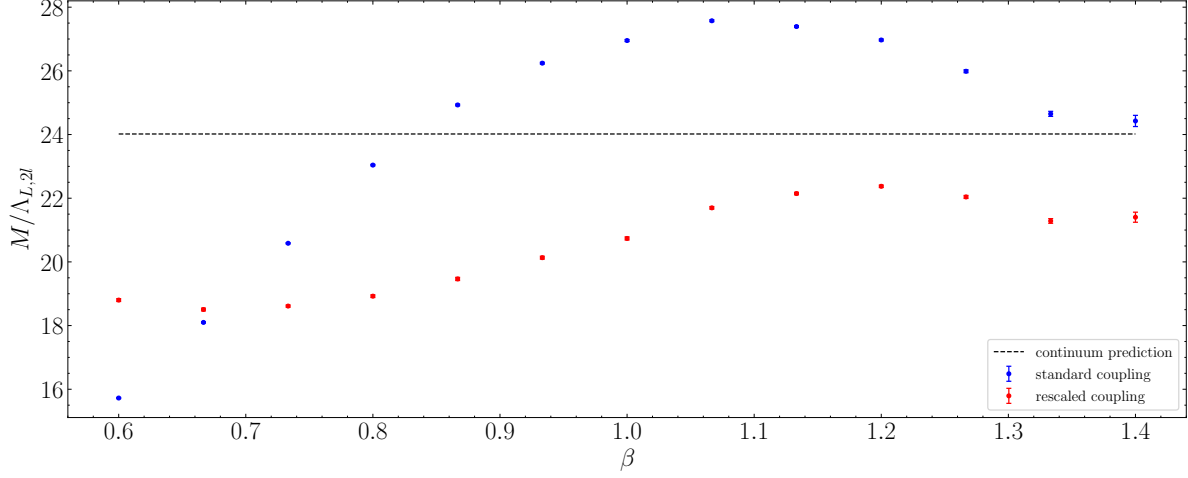


Figure 18. Comparing the mass over Λ ratio using the standard coupling (blue, eq. 143) and the redefinition employed in [5] (red, eq. 165). The same data of tab. 1 is used for both data series.

The argument of the lattice Λ parameter differs by a factor of 4 due to the normalisation of the action (and thus the definition of β) used in this report. The mass over Λ ratio in the new scheme is thus

$$\left. \frac{M}{\Lambda_{2l}} \right|_e = \frac{1}{\xi} \frac{\Lambda_e}{\Lambda_L} \frac{1}{\Lambda_{e,2l}(\beta_e)} = \frac{1}{\xi} e^{\pi \frac{N^2-2}{4N^2}} \frac{e^{8\pi\beta_e}}{\sqrt{8\pi\beta_e}}. \quad (165)$$

In fig. 18, this ratio is plotted against β . For better comparison, the original results from eq. 143, as well as the continuum limit prediction of eq. 144, are shown. The motivation behind this non-perturbative change of variable is to soften the dip in the beta function, which is believed to cause the overshooting for intermediate values of β . While this prescription was successful for the $N = 3, 6, 9, 15$ models studied in [5], we observed no improvement for the $N = 2$ case.