Reduced critical slowing down for statistical physics simulations

Kurt Langfeld
School of Mathematics, University of Leeds, Leeds, LS2 9JT, UK

Pavel Buividovich, P.E.L Rakow, and James Roscoe
Department of Mathematical Sciences, University of Liverpool, Liverpool, L69 7ZX, UK

(Dated: April 12, 2022)

Wang-Landau simulations offer the possibility to integrate explicitly over a collective coordinate and stochastically over the remainder of configuration space. We propose to choose the so-called "slow mode", which is responsible for large autocorrelation times and thus critical slowing down, for collective integration. We study this proposal for the Ising model and the LLR method as simulation algorithm. We firstly show that in a phase with spontaneously broken global symmetry, autocorrelation times grow exponentially with system size for the standard heatbath update. Identifying the magnetisation as collective coordinate, we present evidence that critical slowing down is absent for almost all observables.

I. INTRODUCTION

Stochastic simulations of lattice theories combined with modern computer resources have rapidly evolved to an exceptional theoretical framework enlightening research areas such as Quantum Field Theory [1] and Statistical Physics [2]. Markov Chain Monte Carlo (MCMC) simulations in conjunction with a local update of the degrees of freedom are ubiquitous in the quiver of possibilities.

In MCMC simulations, a bunch of local updates - usually called MC sweep - result into a new configuration of degrees of freedom on the lattice. The simulations generates sequentially a string of lattice configurations. Under the Markov assumption, any configuration only depends on its predecessor. Objects of interests are expectation values. By virtue of the law of large numbers [3], those can be estimated using the N configurations of the Markov set:

$$\langle A\rangle \;\approx\; \frac{1}{N}\sum_{i=1}^{N} A_i \;.$$

The price to pay for a finite reach N is that the above estimator is afflicted by a statistical error ϵ_A, which scales like $1/\sqrt{N}$ under the Markov assumption.

In practical Monte-Carlo simulations, configurations are correlated over a characteristic number of Monte-Carlo updates $t \approx \tau$, which is called auto-correlation time (we will give a proper definition below). An immediate impact is that the statistical error now scales like $\sqrt{\tau/N}$. Large autocorrelations times severely limit the usefulness of simulations at moderate computational costs, and a good deal of algorithmic research has been devoted to simulation methods with small auto-correlations.

The auto-correlation time depends on the simulation algorithm, the parameters of the simulated theory and the system size, say volume V, which could be the number of lattice sites. Of particular interest for many applications is a parameter regime that leaves the lattice degrees of freedom correlated over a typical spatial scale ξ (correlation length). In Solid State Physics, ξ diverges at a second order phase transition. In quantum physics simulations $1/\xi$ acts a regulator for the inherent divergencies of the underpinning quantum field theory, and the limit $\xi \to \infty$ is of crucial importance to extract physics relevant information from those computer simulations. For the important class of *local update algorithms* , there is a monotonically increasing function $\tau(\xi)$ which describes the connection between correlation length ξ and the autocorrelation time τ. On a finite lattice, say with an extent L, spatial correlations are limited by L, leaving us with: $\tau = \tau(L)$. We will distinguish between a power-law and an exponential relation, and refer to the later as *critical slowing down*:

$$\tau(L) \propto L^z \;, \qquad \tau(L) \propto \exp\{m\, L\}.$$

Because of the connection between autocorrelation time τ and statistical error ϵ, theories in the parameter regime afflicted by critical slowing down can only be simulated for finite lattice sizes L, and extrapolation to large L might or might not be possible.

Over many decades, research has been analysing the combination of theories and algorithms studying autocorrelations times for particular observables. For Markov chain simulation that satisfy detailed balance, large autocorrelations times are traced back to low eigenvalues of the transition matrix [4]. The latter paper offers a detailed study for lattice QCD and the important Hybrid Monte Carlo approach [5]. In theories that admit a characterisation of configurations by topology, such as QCD and CP(N) models, critical slowing down is often related to slowly-evolving topological modes [6, 7]. More generally, modes with slowest de-correlation typically correspond to long-wavelength modes of physical fields.

To alleviate the "slow mode relaxation" issue, multigrid methods have been proposed already in the late eighties [8]. So-called *cluster update algorithms* [9, 10] are acclaimed as a solution to the critical slowing down

issue. They are available only for certain models but perform with a small dynamical critical exponent z. Other attempts are a based on a reformulation, and simulations include non-local updates. An example is the so-called worm algorithm [11]. For the CP(N-1) model, which is plagued by the slow mode issue due to topological sectors, a complete absence of critical slowing down was reported in [12] for two dimensions.

Lattice theories that show *spontaneous symmetry breaking* in the infinite volume limit are particularly prone to large autocorrelation times when simulated with local update algorithm in the broken phase. Let ϕ_x be the fields of such a theory with partition function

$$Z(\beta) \;=\; \int \mathcal{D}\phi \; \exp\{\beta\, S(\phi)\} \;,$$

and $M(\phi)$ the order parameter. For any finite lattice size, the symmetry implies that the expectation value of the order parameter, i.e., $\langle M\rangle$ vanishes. In the broken phase, stochastically "important" configurations cluster in domains with $M(\phi)\neq 0$, and $\langle M\rangle$ vanishes upon averaging over these relevant domains. Local update algorithms usually fail to induce transitions between these domains leading to excessively large autocorrelation times.

In this paper, we propose a novel approach to overcome the critical slowing down that is based on the LLR, or density-of-states, approach. Rather than leaving the essential domain average to the MCMC simulation, we calculate the partition function by integrating *explicitly* over the order parameter and stochastically over the remainder of the configuration space. To this aim, we exploit the identity

$$\begin{aligned} Z(\beta) &= \int dm\; \rho(m) \;, \\ \rho(m) &= \int \mathcal{D}\phi\; \delta\Big(m - M(\phi)\Big) \exp\{\beta\, S(\phi)\} \;, \end{aligned}$$

where δ is the Dirac $\delta-$function. Thereby, ρ is called the density-of-states. Density-of-states techniques have seen remarkable successes over the last decade ranging from a study of the QCD phase diagram at significant baryon chemical potentials [13], a recent study of the topological density in pure Yang-Mills theories [14] and the first proof of concept of solving a strong sign-problem using the Z_3 theory [15].

Key to the success of the density-of-states techniques is a robust method to estimate the density-of-states ρ including control over its stochastic errors. In 2012, an efficient approach was proposed to estimate the marginal distribution $\rho(m)$ called density-of-states [16–18]. The so-called LLR-algorithm falls into the class of Wang-Landau methods [19, 20]. Our proof of concept theory will be the 2-dimensional Ising model. We will demonstrate an autocorrelation time that exponentially increases with the system size for the standard heat-bath algorithm, and then present numerical evidence that autocorrelation times for LLR algorithm only increase by a power-law in size. We cannot offer a rigorous proof that critical slowing down is strictly absent, but our numerical findings suggest that the autocorrelation time is strongly reduced for phenomenologically relevant parameters and system sizes.

II. UNDERSTANDING CRITICAL SLOWING DOWN

A. Accessing autocorrelations

The well-studied Ising model also serves here to illustrate the breakdown of *importance sampling* due to failure of ergodicity. The purpose of this section is to quantify this breakdown for the popular Markov-Chain Monte-Carlo (MCMC) approach. We are particularly interested in the parameter dependence of failure, foremost its dependence on the system size. All numerical illustrations of this section are carried out using shockingly small lattice sizes. This illustrates the severeness of the issue: These small sizes are mandatory because of the rapid breakdown of ergodicity at even moderate lattice sizes.

Protagonists are the Ising spins $s_x = \pm 1$ associated with each lattice site x of the lattice of size $V = L\times L$. We use periodic boundary conditions [21] throughout the paper. Partition function Z and action S are given by

$$Z = \sum_{\{s_x\}} \exp\{\beta S\} \;, \qquad S = \sum_{\langle xy\rangle} s_x s_y \;, \tag{1}$$

where the sum in the action extends over all nearest neighbours x and y. Results for autocorrelations will depend on the algorithm. We therefore present details of the simulation here. We are employing the standard *heatbath algorithm* as benchmark:

1. Choose a site x of the lattice at random, and calculate the sum over the neighbouring spins:

$$b_x \;=\; \sum_{y\in\langle xy\rangle} s_y \;.$$

2. Define

$$p_x \;=\; \frac{1}{1+\exp\{-2\beta\, b_x\}} \;,$$

and choose $s_x = 1$ with probability p_x and set $s_x = -1$ otherwise.

3. Repeat both steps 1-2 above V time to complete one lattice sweep.

4. The spin configuration $\{s_x\}_k$ after k sweeps is considered as part of a chain of configurations labeled

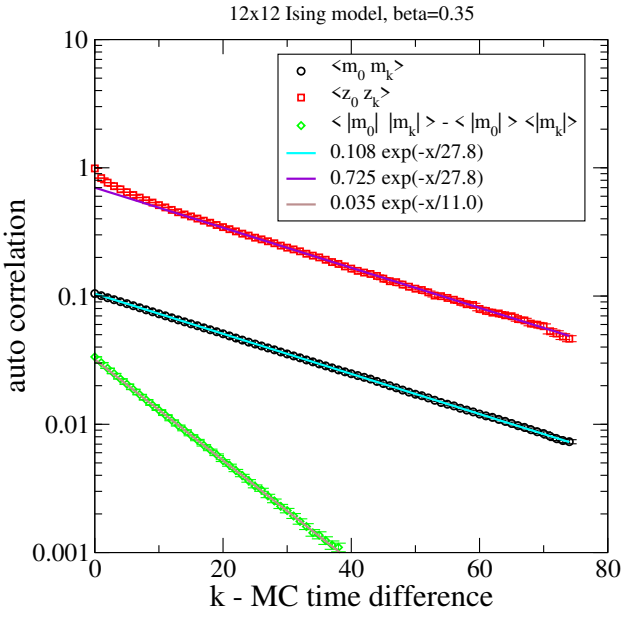


FIG. 1. Auto-correlation functions for a 12×12 Ising model at $\beta = 0.35$ as a function of the MC time difference (right).

by the Monte-Carlo time $k = 1 \ldots N$. Define a sequence of random numbers for an observable $f(\{s_x\})$ by

$$f_1 \to f_2 \to \ldots f_N \,, \qquad f_i \;=\; f\Big(\{s_x\}_i\Big).$$

5. Obtain estimators for observables by

$$f \;:=\; \frac{1}{N}\sum_{i=1}^{N} f_i \,.$$

6. Repeating steps 1-5 many times defines a random process for f itself. We denote the corresponding average by $[f]$. Note that $[f]$ is hence independent of, e.g., the random numbers used for a particular run, but does depend on N. Approximate

$$\langle f\rangle \;\approx\; [f] \,.$$

A variable of particular interest is the *magnetisation per spin*

$$\langle m\rangle \;=\; \left\langle \frac{1}{V}\sum_{x}^{V} s_x \right\rangle \;=\; \langle s_x\rangle \,,$$

which does not depend on the site x due to translation invariance. The corresponding elements of the chain of random variables are given by

$$m_i \;=\; \frac{1}{V}\sum_{x=1}^{V} s_x^{(i)}, \tag{2}$$

where $s_x^{(i)}$ is the spin at site x of the configuration $\{s_x\}_i$.

By the law of large numbers, we find

$$\langle m\rangle \;=\; \lim_{N\to\infty}[m](N) \,.$$

Any stochastic simulation, however, resorts to a finite length N of the chain, and the central question is to what extent is the approximation

$$\langle m\rangle \approx [m] \tag{3}$$

valid?

To avoid a cluttering of notation, we preemptively use a result of the next subsection. By virtue of a symmetry argument, we have

$$\langle s_x\rangle \;=\; 0 \,, \qquad [m](N) \;=\; 0 \,, \quad \forall N \,.$$

As usual, the error for the approximation (3) is given by the standard deviation

$$\epsilon^2 \;=\; \big[m^2\big] \;-\; [m]^2 \;=\; \big[m^2\big] \,. \tag{4}$$

We find

$$\epsilon^2 \;=\; \left[\left(\sum_{i=1}^{N}\sum_{k=1}^{N} m_i\right)\right] \;=\; \frac{1}{N^2}\sum_{i=1}^{N}\sum_{k=1}^{N}[m_i\, m_k] \,. \tag{5}$$

Apparently, the latter equation depends how the random variable m_i is correlated to the variable m_k, and the average $m_i m_k$ is called *auto-correlation*. A key assumption here is that this correlation decreases exponentially with the distance $|i-k|$ between the positions in the chain:

$$[m_i\, m_k] \;=\; m_0^2 \,\exp\left\{-\frac{|i-k|}{\tau}\right\} \,, \tag{6}$$
$$m_0^2 := \big[m_i^2\big] \,,$$

where τ is called *auto-correlation* time. Inserting (6) into (5), the double sum can be performed analytically:

$$\begin{aligned}\epsilon^2 &= \frac{m_0^2}{N^2}\sum_{k=1}^{N}\sum_{i=1}^{N} a^{|i-k|} \\ &= \frac{m_0^2}{N}\,\frac{1+a}{1-a} - \frac{2a\, m_0^2}{N^2(1-a)^2}\,\left(1-a^N\right) \,, \qquad (7)\\ a &= \exp\{-1/\tau\} \,. \qquad (8)\end{aligned}$$

For a moderately sized autocorrelation time, we might find us in a situation where we have $1 \ll \tau \ll N$. Expanding (7) yields for this case:

$$\epsilon^2 \;=\; \frac{2m_0^2\,\tau}{N} \;+\; \mathcal{O}\left(\frac{\tau^2}{N^2}\right) \,. \tag{9}$$

This the famous $1/\sqrt{N}$ law of MCMC simulations taking into account an auto-correlation time $\tau \gg 1$.

I case that the autocorrelation time is exceedingly large, we might face the ordering $1 \ll N \ll \tau$. Expanding (7) for this scenario yields an entirely different picture:

$$\epsilon^2 \;=\; m_0^2 \left[1 - \frac{N}{3\tau} + \mathcal{O}\left(\frac{1}{N\tau}, \frac{N^2}{\tau^2}\right)\right] . \qquad (10)$$

In this case, the error is of order one, and we cannot expect that (3) yields a meaningful approximation. Note, however, that equation (10) still can provide information on the (large) autocorrelation time by virtue of the correction to the leading term.

B. Symmetry breaking and ergodicity

Partition function and action are invariant under a Z_2 transformation of the spins:

$$s_x \;\longrightarrow\; (-1)\, s_x \qquad \text{for} \quad \forall x\, . \qquad (11)$$

This means that the configurations $\{s_x\}$ and $\{-s_x\}$ have the same probabilistic weight implying for any *finite* lattice size V:

$$\langle m\rangle = \langle s_x\rangle = -\langle s_x\rangle = -\langle m\rangle, \;\; \Rightarrow \;\; \langle m\rangle = 0.$$

It also implies that, during the generation of the MCMC chain, the sequence

$$m_1 \to m_2 \to \ldots m_N \quad \text{and} \quad -m_1 \to -m_2 \to \ldots - m_N$$

occur with equal probability , meaning the average over chains vanishes as well, i.e.,

$$[m](N) \;=\; 0\, .$$

The above symmetry enables us to cast each configuration of the MCMC chain into Z_2 classes. To this aim, we define

$$m_i \;=\; z_i\, |m_i| \, , \qquad z_i = \pm 1\, . \qquad (12)$$

Thus, the mapping

$$\{s\}_i \;\longrightarrow\; z_i$$

assigns a Z_2 sector (by virtue of the value of z_i) to each configuration. The symmetry transformation (11) maps each configuration onto a configuration with *equal* statistical weight of the other Z_2 sector.

The above conclusions are not necessarily true in the *infinite volume* limit $V \to \infty$. For infinite systems, the Z_2 symmetry can be *spontaneously broken*. In fact, the Ising model is a prototype to explore this phenomenon. For $\beta > \beta_c$, the statistical system "freezes" in one of the Z_2 sectors with $\langle m\rangle \neq 0$. For $\beta < \beta_c$, we still find $\langle m\rangle = 0$ and the symmetry is realised. The critical value β_c can be calculated analytically [22], and one finds:

$$\beta_c \;=\; \frac{1}{2}\ln\left(1+\sqrt{2}\right) \;\approx\; 0.440686\ldots\, . \qquad (13)$$

This phenomenon is called *spontaneous symmetry breaking* and only applies to infinite volume systems.

Why should we be concerned with this phenomenon since we are only dealing with cases where V is finite? The answer is that most importance sampling algorithms (if not all) for large enough $\beta|\beta_x$ and system size L, anticipate this phenomenon leading to the wrong result

$$[m](N < N_c) \;\neq\; 0$$

even at finite size V. The theorem of large numbers only guarantees $[m] = 0$ for $N \to \infty$, and on some practical applications N_c can be unfeasibly large.

Let us study this statement in the context of an actual numerical simulation. We generate a chain for the magnetisation m_i and for the Z_2 element z_i as a function of the Monte-Carlo time k for $\beta = 0.35$ and $L = 12$. We observe that system changes between Z_2 sectors during the run, which is expected since the Z_2 symmetry is unbroken at such small values of β. However, we realise that regions of positive (negative) m_i cluster for some time. This indicate that we observe a significant autocorrelation time τ even at this small β. In order to quantify this, we present estimators for the auto-correlation functions for

$$m_k, \quad z_k \quad \text{and} \quad |m_k|\, .$$

Note that averages for $[m_k]$ and $[z_k]$ vanish but that for $[|m_k|]$ is non-zero due to the (semi-)positive nature of the observable. The simulation is carried out for a 12×12 lattice at $\beta = 0.35$, which is well placed in the symmetric phase with a moderate auto-correlation time. The simulation starts with a random spin configuration (hot-start) and initially discards 1000 configurations for thermalisation. The result for the auto-correlation functions is shown in figure 1, right panel. Our findings suggest that the auto-correlation functions of m and z are proportional (at least for sufficiently large a MC-time difference), i.e.,

$$[m_i m_k] \;\approx\; m_z^2\, [z_i z_k]\, , \qquad (14)$$

where m_z^2 is a parameter, which can be obtained comparing the fits in figure 1, right panel, and which is about 0.149. This finding signals that the autocorrelation of the centre sector drives the overall autocorrelation of the magnetisation.

We have systematically studied the error ϵ (as given by the equation (4)) for a $L = 12$ lattice size and the three β values 0.3, 0.35 and 0.44. We fitted the theoretical expression for ϵ from (7) (the square root of (7) to be

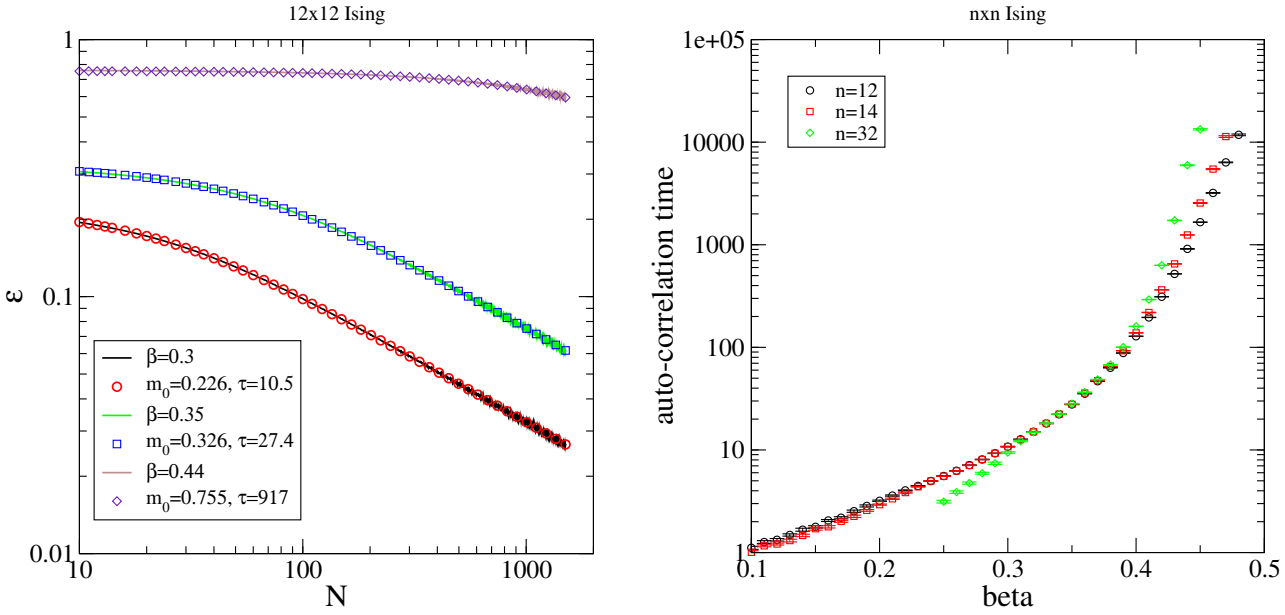


FIG. 2. Left panel Solid lines are estimates (see (17)) for the statistical error ϵ as a function of the length N of the MCMC time series. Open symbols are the theoretical prediction (7).

precise) to the numerical data. This yields an estimate for m_0^2 and the desirable autocorrelation time τ. Our findings are summarised in figure 2, left panel. For β 0.3 and 0.35 the observed autocorrelation time is small enough so that we can observe the characteristic $1/\sqrt{N}$ behaviour at large N. Note, however, that close to $\beta \approx \beta_c$, we observe a large autocorrelation time, which does not allow for the characteristic falloff for the range of N explored. Note, however, that we still can get an estimate for τ by virtue of (7), which does *not* assume $N \gg \tau$.

The same Figure 2, right panel, shows the autocorrelation time as a function of β for the three lattice size 12, 14 and 32. We observe that the autocorrelation time increases *exponentially* in all cases. Note, however, that the slope of the increase changes around $\beta \approx \beta_c$ and is "steeper" for the symmetry broken phase.

Equation (14) suggests that tunneling between Z_2 sector is suppressed and that this suppression is at the heart of the practical ergodicity issue. For each step in of the MCMC chain, we can assign a probability p that the configuration changes the Z_2 sector during this step. We then can calculate the autocorrelation $[z_i z_k]$ analytically.

In a time series of $k+1$ samples z_i, $i = 1 \ldots k+1$ assume that ℓ transitions occur at k possible locations (links between i and $i+1$). The probability for this event is given by

$$\binom{k}{\ell} p^\ell (1-p)^{k-\ell} .$$

The contribution of this event to the auto-correlation function $\langle z_1 z_{k+1} \rangle$ is $(-1)^\ell$. Hence, we find

$$\begin{aligned} \langle z_1 z_{k+1} \rangle &= \sum_\ell^k \binom{k}{\ell} p^\ell (1-p)^{k-\ell} (-1)^\ell \\ &= (1-2p)^k . \end{aligned} \tag{15}$$

Using the latter result in (14) and exploiting the connection to the auto-correlation time in (6), we find the connection between auto-correlation time τ and sector tunneling probability p:

$$p = \frac{1}{2}\left(1 - \mathrm{e}^{-1/\tau}\right) \approx \frac{1}{2\tau} . \tag{16}$$

The latter approximation holds for $\tau \gg 1$. For the example of the previous subsection, i.e., the heat-bath algorithm, a 12×12 lattice and $\beta = 0.35$, we found $\tau \approx 28$ leaving us with a tunneling probability of just $p \approx 1.8\,\%$.

C. Computational resources and precision

The strategy of comparing the performance of two different algorithms is as follows: we will agree at certain level of error ϵ^2 and then ask the question how many "lattice sweeps" N do we need to achieve this.

We already worked out a connection between ϵ^2 and N (see (7)), and it depends on only two parameters, i.e., m_0 and τ. It is time to put this equation to the test. We have generated a time series of $6,000,000$ magnetisations m_k, which we divide into subsequences of length N. For each subsequence, we calculate the average magnetisation

$$m^{(\alpha)} = \frac{1}{N}\sum_{k=1}^{N} m_k^{(\alpha)},$$

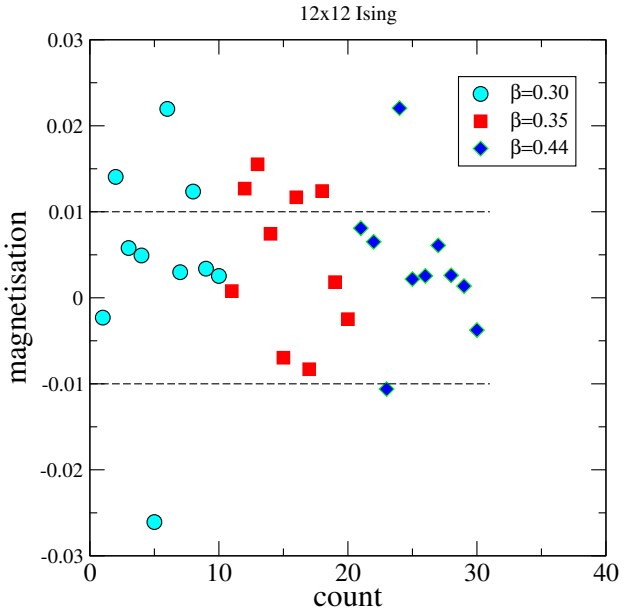


FIG. 3. Average magnetisation from a MCMC time series of length N for three β (see (18) for the β-N pairs).

where α numbers the subsequences from 1 to n_α, which fit into the series of $6,000,000$ magnetisations. The error for the magnetisation estimator (4) is then estimated by

$$\epsilon^2(N) \;\approx\; \frac{1}{n_\alpha}\sum_{\alpha=1}^{n_\alpha}\left[m^{(\alpha)}\right]^2 \;. \qquad (17)$$

Our numerical findings for $N = 10\ldots1500$ appear in figure 2, left panel, as solid lines. We show results for $\beta = 0.3$, $\beta = 0.35$, $\beta = 0.44$. Each curve is fitted by the theoretical prediction (7) with respect to only two fit parameters: m_0 and τ. The agreement is excellent.

We can now ask the question: al least how many MCMC configurations do we need to achieve $\epsilon < 0.01$. For an answer, we use (7) with the readily obtained fit parameter m_0 and τ. The agreement between theory and numerical data is that good that we can extrapolate to N values bigger than 1500. We find that for our lattice size $L = 12$ N has at least to be:

$$\begin{aligned}\beta = 0.30 &: \; N = 10,800 \qquad (18)\\ \beta = 0.35 &: \; N = 58,300\\ \beta = 0.44 &: \; N = 10,460,000\;.\end{aligned}$$

Note that the above N values are vastly outside the fitting range of $N = 10\ldots1500$ and the application of (7) is an extrapolation. It is therefore in order to check the predictions (18). To this aim, we have created, for each β, an MCMC time series of length N and have calculated the corresponding average magnetisation. We have repeated this 10 times. Since $\langle m\rangle = 0$, we expect these m values to be scattered around zero with an error band $\epsilon = 0.01$ (one standard deviation). Our result is shown in figure 3. We observed the expected behaviour even for $\beta = 0.44$, for which $N = 10,460,000$.

It appears that fitting ϵ-data with (7) is an economical way to calculate the auto-correlation time. We have done this for a range of β values and show the result in figure 2, right panel. We observe that the auto-correlation time τ exponentially increases with β. In the "symmetric phase" $\beta \ll 0.44$, the slope seems to be independent of the lattice size L. In the "broken phase" $\beta > 0.44$, the picture changes: the slope of the exponential increase depends on the volume and is significantly bigger than in the symmetric phase. This signals a breakdown of validity of the heat-bath simulation for reasonable sized sample sizes N.

D. Volume dependence and Critical Slowing Down

Of particular interest is to study the volume dependence of the autocorrelation time at give value of β. For subcritical values, i.e., $\beta < \beta_c$, we expect a power-law increase with the system size. This is simply because of that we operate with a local update algorithm, for which it is increasingly difficult to disorder a lattice configuration with increasing size. In the broken phase, i.e., $\beta > \beta_c$, the picture is entirely different: the tunneling between centre-sectors is exponentially suppressed and a changing a Z_2 sector needs resources with exponentially increase with volume. In this subsection, we will verify this picture with unprecedented numerical evidence.

For extracting the autocorrelation time τ for given size L and β, we calculate the autocorrelation function as a function of the Monte-Carlo time t. We fit the asymptotic tail to a the exponential form:

$$C(t) \;=\; [m_0 m_t] \;\propto\; \exp\{-t/\tau\}\;.$$

For small t, we expect power-law corrections to the above functional form and, for large t, the signal might be drowning in the statistical noise of the estimator. Let $E(t)$ be the estimated error of the function $C(t)$ at time t. For the parameters L, β explored in this section, we only take data into with

$$t > 200, \qquad t < t_{\max},\;,$$

where

$$t_{\max}: \quad \text{largest } t \text{ with: } C(t) > 5E(t)$$

or $t_{\max} = 2000$ whatever is smaller. This is necessary to keep memory usage under control during the simulation. One of our many results is shown in figure 4, top panel. Parameters have been $L = 16, 32$ and $\beta = 0.43$. Not all data are shown since the figure would become too crowded. The numerical data is well fitted by exponential form. Throughout, we monitor the χ^2 of the fit. Errors for the fit parameter and hence the autocorrelation time

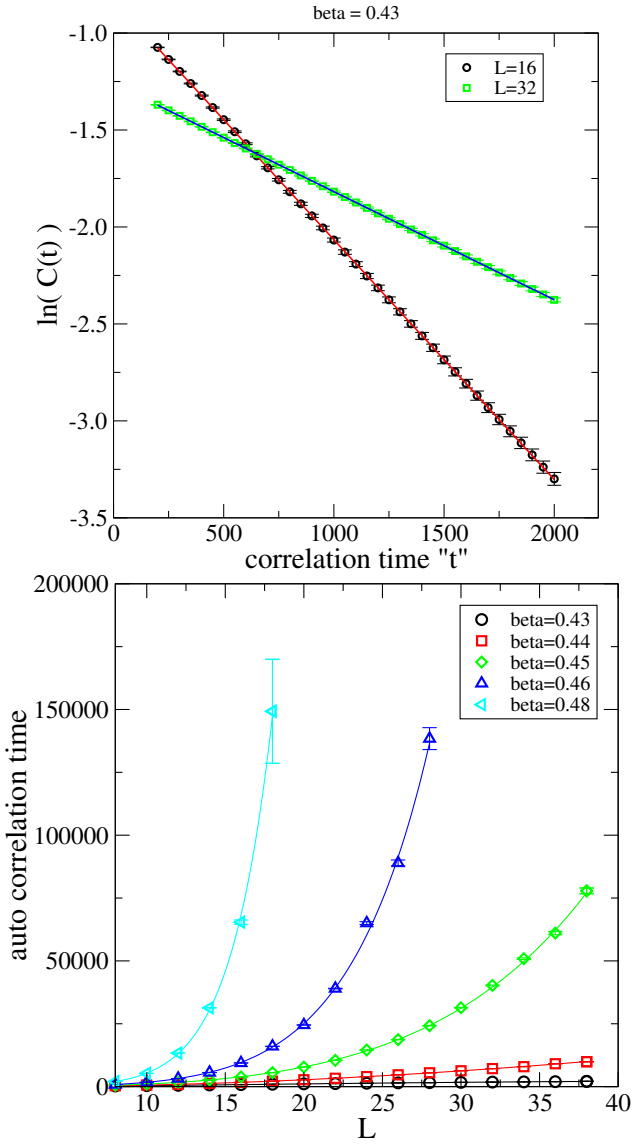


FIG. 4. Autocorrelation function as a function of Monte-Carlo time t for two lattice sizes at $\beta = 0.43$ (top). The autocorrelation time for the magnetisation as observable as a function of system size L for several values of β (bottom).

is obtained by bootstrap. For the fits shown in figure 4, we obtained specifically

$$\tau(L=16) \;=\; 808.5(6)\;, \qquad \tau(L=32) \;=\; 1794(1)\;.$$

We have repeated this analysis for $L \in [8, 39]$ and $\beta = 0.43, 0.44, 0.45, 0.46, 0.48$. The results for the autocorrelation time τ is shown in the same figure 4, bottom panel. We observe that τ rapidly grows for β values instigating spontaneous symmetry breaking. We observe that the numerical data for τ are well fitted by the formula

$$\tau(L) \;=\; b_0\, L^{b_1}\; \exp\{b_2\, L\}\;. \tag{19}$$

In the absence of the exponential ($b_2 = 0$), the formula describes a power-law growth of τ with size L while, for $b_2 > 0$, the formula suggests an dominating exponential growth. The fits are also shown in the bottom panel of figure 4. They well describe the data. In particular, we find:

	$\ln(b_0)$	b_1	b_2
$\beta = 0.43$	1.727(8)	1.991(4)	−0.0035(2)
$\beta = 0.44$	1.213(7)	2.303(3)	−0.0087(1)
$\beta = 0.45$	1.26(3)	2.26(2)	0.047(1)
$\beta = 0.46$	1.5(1)	2.00(7)	0.130(4)
$\beta = 0.48$	1.5(10)	1.8(7)	0.28(6)

TABLE I. Results of the fitting of the lattice size dependence of the autocorrelation time in Monte-Carlo simulations with heatbath updates with a product of power law and exponential functions (19).

We thus find evidence that b_2 starts growing to non-zero values around the critical values $\beta \approx \beta_c$ for the phase transition. In the symmetric phase at $\beta = 0.43$, we find that the autocorrelation time τ approximately grows with the volume L^2.

III. REDUCED CRITICAL SLOWING DOWN WITH THE LLR METHOD

A. Brief introduction to the LLR approach

We are aiming to estimate the magnetisation M with reliable errors over a wide spectrum of β-values stretching from the symmetric phase deep into the symmetry broken phase for $\beta \gg 0.44$. We start by defining the density-of-states $\rho(M)$ for the magnetisation:

$$\rho(M) \;=\; \frac{1}{Z} \sum_{\{s_x\}} \delta\Big(M, \sum_x s_x\Big)\; \exp\{\beta S\} \tag{20}$$

with the action S in (1). The Kronecker delta is defined in the usual way:

$$\delta(i,k) = 1 \qquad \text{for } i = k\;, \qquad 0 \text{ else.}$$

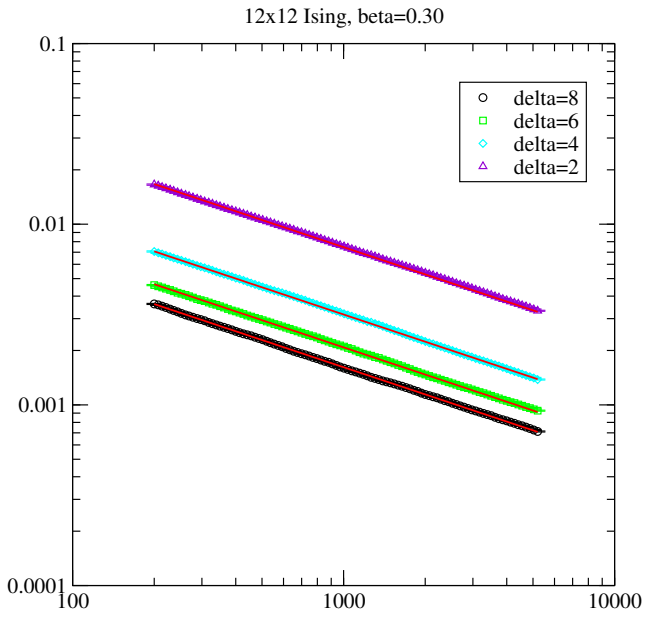
The magnetisation is then given by

$$\langle m \rangle \;=\; \frac{\sum_M M\; \rho(M)}{\sum_M \rho(M)}\;,$$
$$M = -V, -V+2, \ldots, V-2, V. \tag{21}$$

With the normalisation

$$\sum_M \rho(M) \;=\; 1 \tag{22}$$

because of the definition (14) and that of the partition function Z in (1), $\rho(M)$ can be interpreted as the probability with which magnetisations M contribute to expectation values such as the one in (21). By virtue of the Z_2



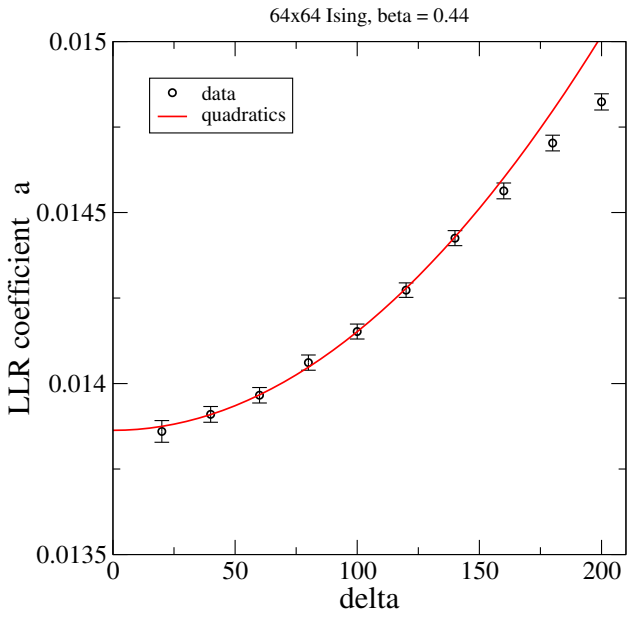


FIG. 5. Left: The error in the LLR coefficient a as a function of the number of Robbins-Monroe iterations n (29). The fits correspond to a $1/\sqrt{n}$ power law. Right: Dependence of the LLR coefficient a on δ for a 64×64 lattice near criticality ($\beta = 0.44$).

symmetry transformation (11), the density is symmetric, i.e.,

$$\rho(-M) \;=\; \rho(M)\;,$$

leading to $\langle m\rangle = 0$ as expected. In our numerical study we will *not* exploit the above symmetry relation but rather will study the stochastic errors for our estimate for $\langle m\rangle$.

At the heart of the LLR approach is the expectation value

$$\langle\!\langle f\rangle\!\rangle(a) = \frac{1}{\mathcal{N}} \sum_{\{s\}} f(s)\; \mathrm{e}^{\beta S + a\, m(s)}\; W_\delta\Big(m_0, m(s)\Big) \qquad (23)$$
$$m(s) = \sum_x s_x\;,$$

where we here use a Heaviside function for the window function:

$$W_\delta\Big(m_0, m(s)\Big) \;=\; \begin{cases} 1 & \text{for } m_0 - \delta \le m(s) \le m_0 + \delta\;. \\ 0 & \text{else.} \end{cases} \qquad (24)$$

Note that $\langle\!\langle f\rangle\!\rangle(a)$ depends also on the parameters δ and m_0, and a is also called the LLR coefficient. You can obtain the density-of-states $\rho(m_0)$ by carrying out the following steps:

1. For a given δ and m_0, solve the stochastic equation

$$\langle\!\langle m(s) - m_0\rangle\!\rangle(a^*) \;=\; 0 \qquad (25)$$

for a (solution a^*), which depends smoothly on m_0 and δ for $m_0 \in [-V, V]$.

2. Use

$$\frac{d}{dm_0} \ln \rho(m_0) \;=\; \lim_{\delta\to 0} a(\delta, m_0) \qquad (26)$$

and evaluate (or estimate) $\rho(m_0)$ up to a multiplicative factor by integrating the above equation.

3. Determine the multiplicative factor by normalising ρ (see (22)).

The last step might be optional since a normalisation constant drop out of expectation values such as the one in (20).

As for the heat-bath MCMC approach, we are interested in the question: what type of precision can we achieve as a function of the invested computational resources. We therefore will critically investigate the parameter dependence of the numerical error.

Let us first comment on solving the stochastic equation of the type (25). This task has been extensively studied firstly by Robbins and Monroe [23] and then taken up by number of authors (see [24] for a review). If $F(a)$ is a noisy estimator for

$$f(a) := \langle\!\langle m(s) - m_0\rangle\!\rangle(a) \qquad (27)$$

Robbins and Monroe propose an under-relaxed iterative approach. Starting with some a_1, consider the recursion

$$a_{n+1} = a_n \;-\; \alpha_n\, F(a_n) \qquad (28)$$

with a sequence of positive weights α_n, $n = 1, 2, 3\ldots$ satisfying

$$\sum_{n=1}^{\infty} \alpha_n \to \infty\;, \qquad \sum_{n=1}^{\infty} \alpha_n^2 \to \text{finite}\;.$$

The sequence converges with probability one to the solution $a^* := a_\infty$ [25]. A particular sequence was suggested by Robbins and Monro:

$$\alpha_n = \frac{\kappa}{n} \; .$$

The algorithm reaches *asymptotically* the optimal convergence rate of $1/\sqrt{n}$, but the initial (low n) performance crucially depends on the sequence. Chung [26] and Fabian [27] showed that optimal convergence is reached with the choice:

$$\alpha_n = \frac{1}{f'(a^*)\; n} \; .$$

This choice, however, hinges on the solution a^*. For the specific problem at hand, i.e., (25), we can, however, find a good value κ. For small enough δ, the marginal for the magnetisation m in the window $[m_0-\delta, m_0+\delta]$ is Poisson distributed, i.e., $\propto \exp\{-a^* m\}$. Together with the 're-weighting' factor $\exp\{am\}$ in (23), the m distribution becomes flat for values m inside the window. We then find with (27), the definition (23) and the solution (25):

$$\begin{aligned} f'(a^*) &= \langle\!\langle (m(s)-m_0) m(s) \rangle\!\rangle (a^*) \\ &= \langle\!\langle (m(s)-m_0)^2 \rangle\!\rangle (a^*) = \frac{1}{2\delta+1} \sum_{m=-\delta}^{\delta} m^2 \\ &= \frac{\delta\,(\delta+1)}{3} \;\approx\; \frac{\delta^2}{3} \; . \end{aligned}$$

The latter hold for $\delta \gg 1$, which would also be the result if the degrees of freedoms have a continuous domain of support. Note that by the nature of the task at hand (25,23), $f'(a^*)$ does not depend on the solution a^*. We arrive at the iteration that we will study in the remainder of the paper:

$$a_{n+1} = a_n \; - \; \frac{3}{\delta^2\; n}\; F(a_n) \; . \tag{29}$$

We put the above iteration to the test for a $V = 12 \times 12$ lattice, $\beta = 0.3$, $m_0 = INT(0.8V)$ and several δ values. The estimator $F(a)$ is obtained by 20 successive lattice sweeps. Our findings for the error ϵ_a in the LLR coefficient a as a function of the Robbins Monro iteration time n is shown in figure 5. We performed $1,000$ independent Robbins Monro runs to estimate the error for ϵ_a. We find optimal convergence behaviour already for $n > 200$. The error for small δ are smaller than those for large δ. This is expected since for larger δ the window function is wider and hence includes more spin in the averaging.

B. Precision versus resource

The following study is done for the 2D Ising model on a 32×32 lattice. The objective is to find the amount of 'lattice sweeps' is needed to calculate the magnetisation $\langle m \rangle$ to a given accuracy. In the last section, we saw that the heat bath algorithm needs a rapidly increasing amount of resource if β approaches the regime of a spontaneously broken symmetry.

Our simulations parameters are "ball park" figures and are *not* fine tuned.

1. We use a step function as window function $m \in [m_0 - \delta, m_0 + \delta]$ with $\delta = 8, 16, 24, 32$.
2. We perform $10,000$ Robbins Monro iterations for each m_0 and for each δ leaving us with an estimate for the LLR parameter $a(\delta)$. We perform a quadratic fit for extrapolating to $\delta \to 0$ and set: $a = a(0)$.
3. Each double expectation value is estimated with 20 lattice sweeps.
4. We generate LLR parameters a for 63 values of m_0, i.e., $(m_0)_k = -32^2 \; + \; k \; \times \; 32$, $k = 1 \ldots 63$.
5. For each m_0, we generate 80 potential LLR parameters a_i for the subsequent statistical analysis.

We will measure resource in units of 'lattice sweeps' (ls), i.e., one resource unit corresponds to V spin updates. This choice allows to measure resource independent of hardware employed for the calculations. All algorithms studied here - heat bath update, cluster algorithms, LLR method - uses 'lattice sweeps' at low level of the calculation. Although Ising spin updates are low cost, the 'lattice sweep' might be the most expensive computational element for other systems such as gauge theories with fermions (QCD) where a lattice sweep could be defined by a Hybrid Monte-Carlo trajectory.

To generate the above data set for the LLR coefficients (steps 1-4), the resources needed are

$$4 \; \times \; 20 \; \times \; 10,000 \times 63\, ls \; = \; 5.04 \cdot 10^7\, ls \; . \tag{30}$$

From this data set, we can already estimate expectations values of functions of the magnetisation, and the objective here is to estimate the precision with which we can calculate $\langle m \rangle$ (which equals zero for a simulation with infinite resources). To this aim, we will repeat the calculation 80 times. This, the analysis uses the resources of $5.04 \cdot 10^7 \times 80\, ls = 4.032 \cdot 10^9\, ls$, which must not be confused with resource (30) needed to produce one sample result.

The density of states ρ for the magnetisation m is obtained by integration of the LLR- coefficient:

$$\rho(m) \; = \; \exp\left\{ \int_0^m a(m')\; dm' \right\} \; . \tag{31}$$

The normalisation is arbitrarily chosen to be $\rho(0) = 1$. Expectations values are then obtained by a second integration, e.g.,

$$\langle m \rangle \; = \; \int m\; \rho(m)\; dm \; / \; \int \rho(m)\; dm \; . \tag{32}$$

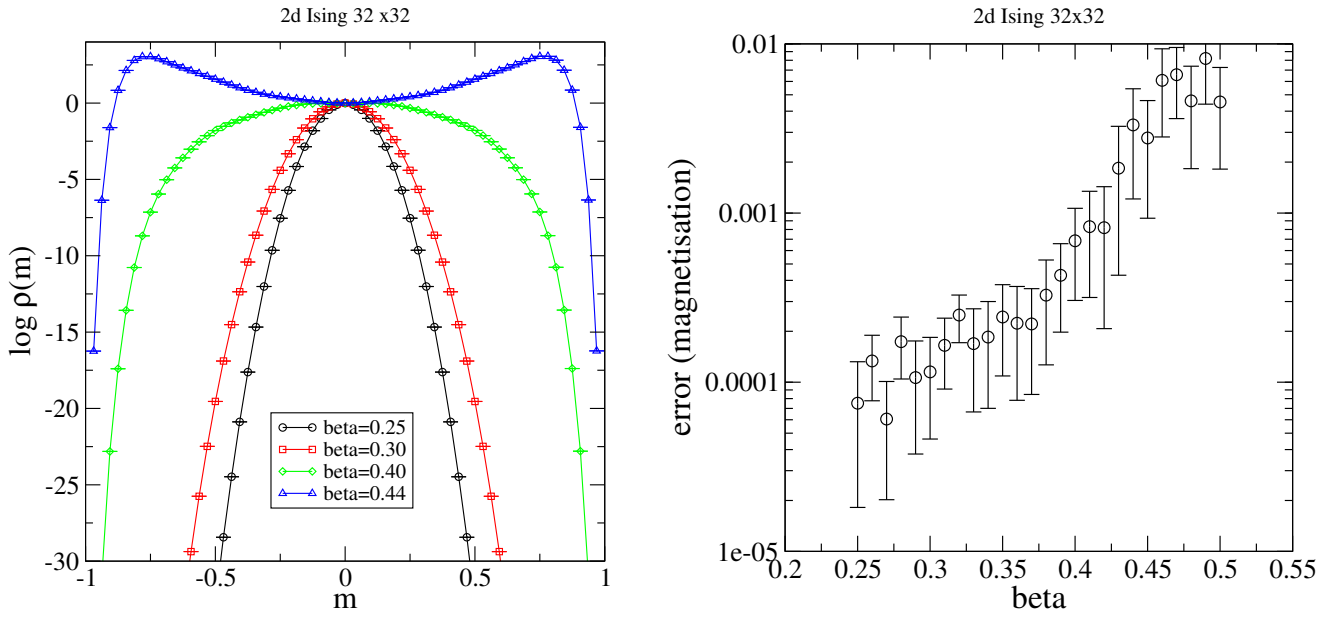


FIG. 6. Left: Log of the density of states $\rho(m)$ as a function of the (intensive) magnetisation m for four β values. Right: Error of the magnetisation (32) as a function of β.

Early studies [15–17] used a trapezium rule and summation, which leads to an accumulation of error for increasing m. Representing the function $a(m)$ by high degree polynomial and performing the integrations (semi-) analytically has proven very successful [18, 28–30]. One can prove that the density of states for Ising model is an even function in m by virtue of its Z_2 symmetry. Correspondingly, the LLR coefficient $a(m)$ is an odd function. A numerical approach exploiting this observation would approximate $a(m \geq 0)$ by polynomial of odd powers of m. This would lead to the exact result $\langle m \rangle = 0$.

The prime objectives here is to avoid any assumptions on symmetry and to observe to what extent the exact result $\langle m \rangle = 0$ is obtained. For this purpose, we approximate $a(m)$ over the full domain by polynomial containing even and odd powers of m. We find that a polynomial of degree 16 represents the numerical data for a very well.

The result for $\rho(m)$ (on a logarithmic scale) is shown in figure 6. Error bars are obtained by the bootstrap method:

1. For each m_0, calculate a set of n_B LLR coefficients from independent runs. We have chosen here $n_B = 60$.

2. For each of the (discrete) m_0 choose an LLR coefficient out of the n_B possibilities.

3. Fit a polynomial of degree 16 to the data.

4. Perform the integration (31) analytically and obtain one sample for $\rho(m)$.

5a. Repeat this procedure many times and calculate the average for $\rho(m)$ and the standard deviation (error bar).

Step 5a gives rise to the graphs in figure 6, left panel. We find that for $\beta = 0.25, 0.30, 0.40$ the density-of states is maximal at $m = 0$ making $m = 0$ the most likely magnetisation. We also observe that, for a finite $L = 32$ lattice, the curve for $\beta = 0.44$ develops a double peak structure, which is characteristic for the spontaneous breakdown of symmetry. We expect that for increasing lattice size, the β for which the double peak structure occurs will approach β_c in (13).

We are here not primarily interested in the density of states ρ but the expectation value of the magnetisation

$$m \;=\; M/V \;=\; \frac{1}{V}\sum_x z_x \;.$$

In this case, we replace step 5a by:

5b. For the sample $\rho(m)$, calculate the two integrals in (31) analytical and, thus, obtain a sample value for $\langle m \rangle$. Repeat this procedure many times and calculate the average for $\langle m \rangle$ and the standard deviation (error bar).

Figure 6, left panel, shows the (log of the) density of states as a function of the intrinsic magnetisation $m = M/L^2$. For the finite volume $L = 32$, we see that the most likely magnetisations m are at $m \neq 0$ for $\beta = 0.44$. This is a precursor of spontaneous symmetry breaking. Increasing the volume, it is expected that this bifurcation moves up in β to approach β_c (13) in the infinite volume limit.

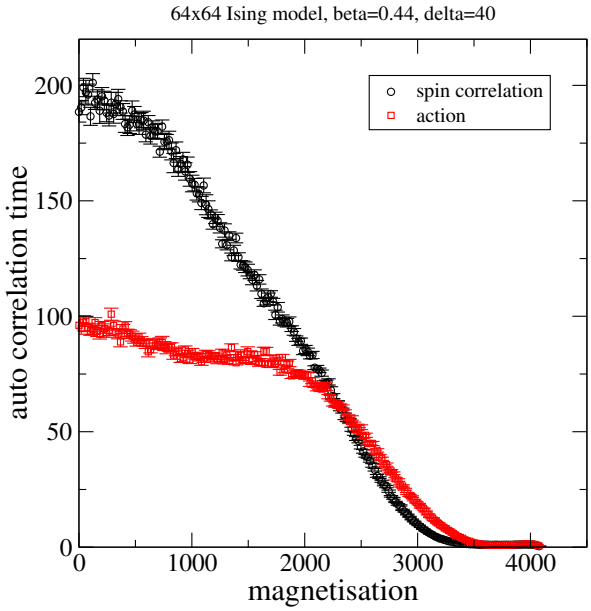


FIG. 7. Autocorrelation time for the LLR double expectation value for the action and the spin-spin correlation as a function of m_0.

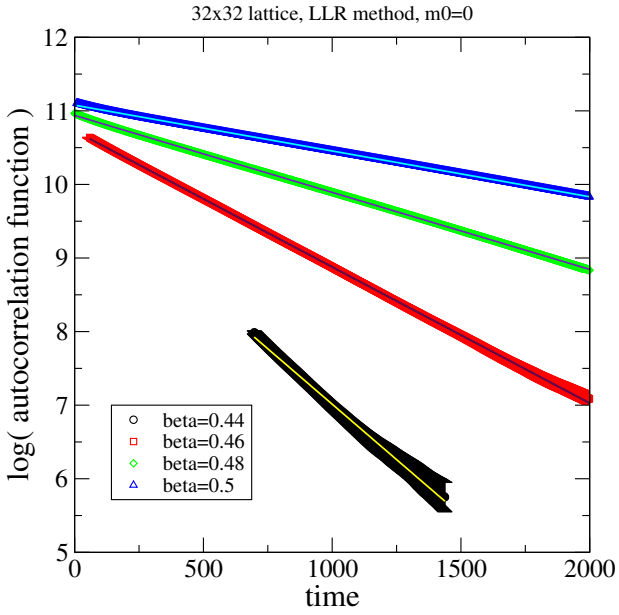


FIG. 8. Autocorrelation function for the observable M_1 (34) as a function of the autocorrelation time for four values of β.

Having calculated the density of states, we estimated the magnetisation m using (32). The precision with which the exact result $\langle m\rangle = 0$ is recovered depends on the quality of the symmetry $\rho(m) = \rho(-m)$. Our result for the *error* of m is shown in figure 6, right panel, as a function of β, where we have kept fixed the number of Robbins Monro iterations and the bootstrap copies. We find a moderate increase with increasing β, which can be explained by the larger variation of $\rho(m)$ with m due its peak structure, which makes it harder to control the numerical precision of the integration over m in the integrals of (23).

C. Autocorrelations and density-of-states

The so-called double expectation values such as in (17) are at the heart of the LLR approach since they ultimately give rise to a and hence the density of states (see (25). These expectation values can be viewed as ordinary Monte-Carlo expectation values, and, as such, they are susceptible to autocorrelations of the Markov chain.

We already established that there is a close link between spontaneous symmetry breaking and the exploding autocorrelation time for local update algorithms operating close to criticality. We expect that the double expectation values are much less affected by this phenomenon simply because they are not operating a close to criticality "most of the time".

We first note that the double expectation values depend on a number of parameters, which are not present in a standard heat bath simulation. There is the LLR parameter a which adds a term $a\sum_x s_x$ to the action. For $a \neq 0$ this parameter acts like a magnetic field, which breaks the Z_2 symmetry $s_x \to -s_x$. Secondly, the window function $W(m_0, m(s)$ (18) is part of the probabilistic measure. It restricts spin configurations to values of the magnetisation $m(s)$ close m_0. This means that this factor also breaks the Z_2 symmetry as long as $m_0 \neq 0$. Note, however, that for $m_0 = 0$, the solution of the stochastic equation is $a = 0$ precisely because of the Z_2 symmetry. We thus expect that the calculation of $\rho(m \approx 0)$ might be affected by long autocorrelations.

Note that for most of the observables in the broken phase, $\rho(m \approx 0)$ might be an entirely suppressed domain of integration for the integrals in e.g. (32). In this case, these autocorrelations have little impact on the precision of the calculation.

In a first step, we studied the autocorrelation time for the action and the spin-spin correlation function for different values of m_0, the centre of the window function:

$$\text{action: } \sum_{\langle xy\rangle} s_x s_y \ , \qquad \text{spin-spin: } s_x\, s_{x+L/2} \ .$$

Our findings are summarised in figure 7, left panel. Indeed, we observe that those auto-correlations are highest close to $m_0 = 0$ where the system can have critical behaviour.

Since the magnetisation is constrained to a region around m_0 in the LLR simulation, autocorrelations of the magnetisation are indeed very small. In search of an observable susceptible to longest autocorrelations, we

introduce the Fourier transform of the magnetisation:

$$\bar{M}(p_x, p_y) \;=\; \sum_x s_{x,y} \;\cos\left(\frac{2\pi}{L}\;(x\,p_x + y\,p_y)\right)\;, \quad (33)$$

For $p_x = 0, p_y = 0$, this quantity becomes the magnetisation, i.e., $M = \bar{M}(0)$. Another "infrared" observable, similarly prone to autocorrelations but unconstrained by the LLR approach, is $\bar{M}$ for the lowest momenta with either $p_x = 1$, $p_y = 0$ or $p_x = 0$, $p_y = 1$. The choice of these observables is motivated by the common observation that low-momentum modes typically have the slowest relaxation/decorrelation rate in local, translationally invariant quantum field theories. We thus study the autocorrelation time for the observable

$$M_1 \equiv \bar{M}(1,0) \;=\; \sum_{x,y} s_{x,y} \;\cos\left(\frac{2\pi}{L}\,x\right) \quad (34)$$

and the analogous quantities with the sin function, and with y as the Fourier variable. To this end, we firstly estimate the autocorrelation function $C(t)$ of M_1 and extract the autocorrelation time by analysing the exponential decrease at large values of t. If t is too large, statistical noise drowns the signal. If $\sigma(t)$ is the standard deviation of the estimator for $C(t)$, we only use data with

$$C(t) \;>\; 5\,\sigma(t)\;.$$

At small values of t, $C(t)$ is not well represented by an exponential function, which only hold asymptotically. We proceed as follows: starting at $t = t_0 = 0$, we fit an exponential function to the data and obtain the χ^2/dof. We then systematically increase t_0 until χ^2/dof falls below 0.8 for the first time. We thus extract the autocorrelation time τ from the fit:

$$a_0 \;\exp\{-t/\tau\}\;.$$

Figure 8 shows the correlations function $C(t)$ for a 32^2 lattice and for four values of β within the dynamically generated domain of support. Repeating this procedure for lattice sizes between $L = 8$ and 48, we find the result shown in figure 9. We indeed observe that the autocorrelation times for M_1 increase with increasing lattice size L, but not nearly to the extent as we have seen those for the heatbath simulation and the magnetisation M.

The central question is whether or not these autocorrelations times increase *exponentially* with L. In search of an answer, we have employed the same fit (19) of the data as in the case of the heatbath result. Of particular interest is the coefficient b_2, which indicates critical slowing down for $b_2 > 0$. our findings are summarised in the table below:

We observe a very small coefficient b_2 when compared to the heatbath simulation where $b_2 \approx 0.28$ at $\beta = 0.48$. The quality are less convincing especially for $\beta = 0.5$. Here, figure 9 shows two fits: the exp-powerlaw

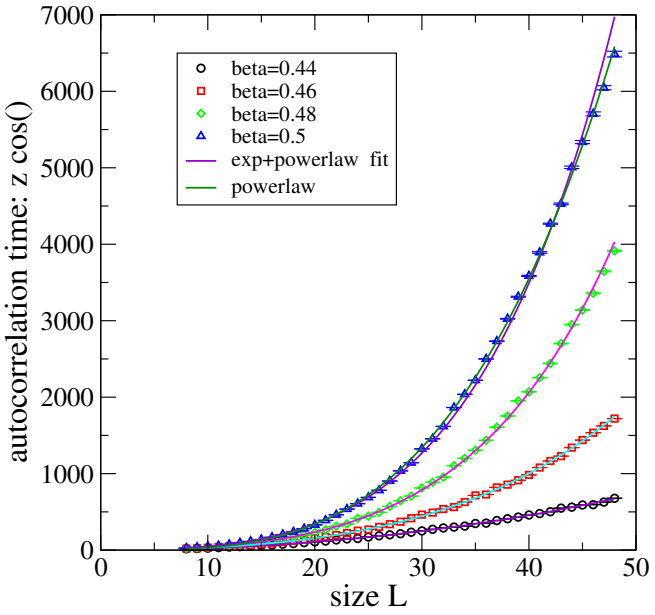


FIG. 9. Autocorrelation time for the observable M_1 (34) as a function of the system size L for four values of β and for the worst case scenario $m_0 = 0$.

	$\ln(b_0)$	b_1	b_2
$\beta = 0.44$	$-1.28(1)$	1.965(4)	0.0038(2)
$\beta = 0.46$	$-1.26(3)$	1.942(3)	0.025(1)
$\beta = 0.48$	$-1.495(6)$	2.080(3)	0.036(1)
$\beta = 0.50$	$-2.194(7)$	2.484(4)	0.029(2)

TABLE II. Results of the fitting of the lattice size dependence of the autocorrelation time in LLR simulations with a product of power law and exponential functions (19).

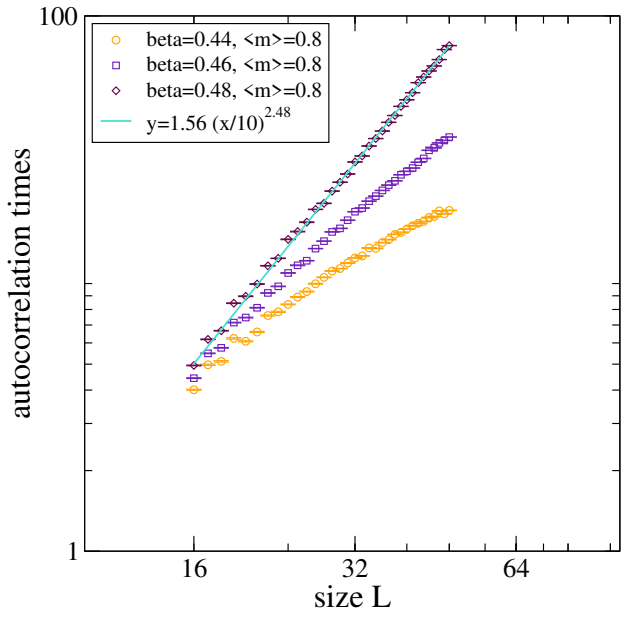


FIG. 10. A comparison of the dependence of autocorrelation time for the observable M_1 on lattice size L for the LLR approach for $m_0 = 0.8$ for several values of β.

fit (19)and a power-law fit $b_2 = 0$. Both fits reasonable well present the data. We are carefully optimistic that any exponential growth is a quite small rate implying that auto-correlation times are manageable for realistic lattice sizes. Higher precision data and perhaps larger lattice sizes are needed to evidence this at a quantitative level.

As detailed above, only the double-expectation values for $m_0 = 0$ are afflicted by criticality since, for $m_0 \neq 0$, the Z_2 symmetry is explicitly broken by the window function and an LLR-coefficient $a \neq 0$. Nevertheless, it is important how the autocorrelation times scale with the lattice size L. In the broken phase, say for $\beta > 0.45$, the marginal distribution for the magnetisations peak at rather large values $M/V \approx \pm 0.9$. For generic observables with a broad domain of support from large portions of the domain of magnetisation, the dominant contributions from the LLR integration over the magnetisation raises from the region around $M/V \approx \pm 0.9$. Hence, we studied the volume dependence of the observable (34) as a function of the lattice size L at $m_0 \neq 0$. The results for $m_0 = 0.9$ are shown in figure 10 in the double-log scale in comparison with the $m_0 = 0$ data. We observe that auto correlation times are orders of magnitudes smaller than in the $m_0 = 0$ case. Most importantly however, we find that the increase of the autocorrelation time with size is at most polynomial in L and for β values away from its critical value even sub-polynomial. Log-log scale plot illustrates this in a particularly clear way, mapping any power-law dependence to a straight line. Therefore plots of functions that grow faster than a power of L appear as bending upwards from a straight line, whereas plots of functions with sub-polynomial growth are bending down from a straight line.

This is an important finding since observables that receive their dominant contribution from the regions of large magnetisation *are not* affected by critical slowing down.

IV. DISCUSSION AND CONCLUSIONS

In this work we used a simple two-dimensional Ising model to demonstrate the potential of LLR algorithm to reduce critical slowing down of Monte-Carlo simulations in a common situation when high potential barriers between different parts of the configuration space make Monte-Carlo updates non-ergodic. While for the Ising model there are efficient model-specific cluster algorithms that eliminate this problem, the advantage of the LLR method is that it is applicable to any lattice field theory. Correspondingly, our analysis is performed without making any explicit assumptions on the symmetry of the model.

Our basic idea is to identify the long-wavelength mode or observable that exhibits the largest autocorrelation time in a Monte-Carlo process. We then decompose the configurations space into the mode and the hyperspace orthogonal to this mode. The LLR methodology allows to integrate the slow mode explicitly while the integration over the hyperspace is done stochastically using MCMC techniques.

For the Ising model, the mode that exhibits the longest autocorrelation time is the global magnetisation, that is, the sum of all spins. We expect that for all models that are well described by the Landau theory of phase transitions the global order parameter will always have the longest autocorrelation time. Our approach also resembles, to some extent, lattice QCD simulations in fixed topological sectors [7]. Indeed, global topological charge is known to be the observable with longest autocorrelation time in lattice QCD.

We also found that once we separate out the order parameter and keep it within an interval of finite width during Monte-Carlo updates, the autocorrelation time for other modes with lowest nonzero momentum $p = \frac{2\pi}{L}$, where L is the linear system size, becomes somewhat larger. In unrestricted Monte-Carlo simulations the autocorrelation time of these modes grows not faster than a polynomial, and only the zero-momentum mode that corresponds to the order parameter exhibits exponential growth of autocorrelation time. In simulations with restricted order parameter, however, the autocorrelation time of these observables can still grow faster than polynomially, as one can see from Fig. 11. The maximal faster-than-polynomial growth is observed for magnetization $m \approx$, whereas in the region of most probable m values the growth is clearly sub-polynomial. Even for simulations with $m \approx 0$, the results of mixed power law times exponential fits in Tables II and I suggest that the coefficient b_2 in front of the lattice size in the exponential factor $\exp(b_2 L)$ is considerably smaller than for the conventional Monte-Carlo simulations. We should stress that the LLR algorithm essentially uses the same heat-bath updates as the conventional Monte-Carlo simulations. The only difference is the presence of the so-called window function and the LLR-parameter a, which couples to the order parameter. Except for the case where the window function restricts configuration to a narrow window around $M_0 = 0$, the window function and $a \neq 0$ implies that the Z_2 symmetry is explicitly broken and the criticality is absent. Indeed, we observe that critical slowing down is absent for magnetisation $M_0 \neq 0$.

We conclude that the LLR algorithm has a potential for solving the issue of critical slowing down for most observables. Only observables that are sensitive to the marginal distribution around $M \approx 0$, no matter how small it is, might be affected by critical slowing down. We only know one such observable: the order-disorder interface tension.

As a next step, it would be interesting to check whether explicit integration over more than one observable using higher-dimensional generalisation of the LLR algorithm

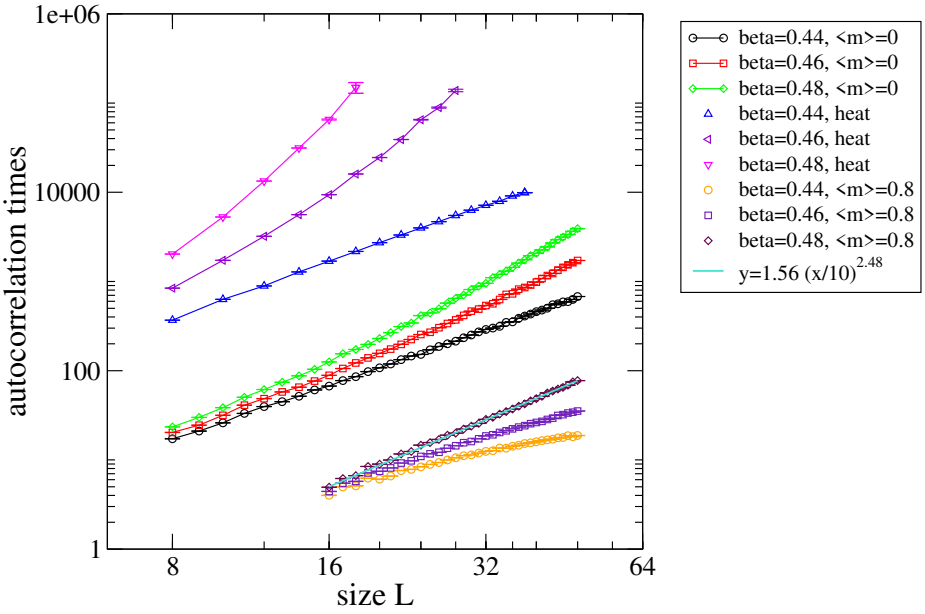


FIG. 11. A comparison of the dependence of autocorrelation time on lattice size L for the conventional heatbath algorithm, where total magnetisation has the longest autocorrelation time, and for the LLR algorithm with magnetisation in the vicinity of $m_0 = 0$ and $m_0 = 0.8$, where the Fourier component of magnetisation with lowest nonzero momentum exhibits slowest de-correlation.

could result in further reduction of computational time. It is also worth exploring whether the application of LLR method to fermionic systems could reduce ergodicity issues related to zeroes of fermionic determinant. Finally, in a recent paper [31] it was suggested that normalising flows can eliminate the need to integrate the density of states over m altogether, thus yielding an even larger speed-up for Monte-Carlo simulations. It would be interesting to see to what extent normalizing flows can further reduce the critical slowing down in our situation.

ACKNOWLEDGMENTS

The numerical simulations were undertaken on ARC4, part of the High Performance Computing facilities at the University of Leeds, UK. PELR is supported in part by the STFC under contract ST/G00062X/1.

[1] H. J. Rothe, *Lattice Gauge Theories*, 4th ed. (WORLD SCIENTIFIC, 2012) https://www.worldscientific.com/doi/pdf/10.1142/8229.
[2] K. Binder and D. Heermann, *Monte Carlo simulation in statistical physics*, 4th ed., 80 (Springer, Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Tokyo, 2002) pp. XII, 180 S.
[3] F.M.Dekking, *A Modern Introduction to Probability and Statistics Understanding Why and How* (Springer, Londan, 2005).
[4] S. Schaefer, R. Sommer, and F. Virotta (ALPHA), Nucl. Phys. B **845**, 93 (2011), arXiv:1009.5228 [hep-lat].
[5] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Phys. Lett. B **195**, 216 (1987).
[6] C. Bonati and M. D'Elia, Phys. Rev. E **98**, 013308 (2018), arXiv:1709.10034 [hep-lat].
[7] R. Brower, S. Chandrasekharan, J. W. Negele, and U. J. Wiese, Phys. Lett. B **560**, 64 (2003), arXiv:hep-lat/0302005.
[8] E. P. Stoll, J. Phys.: Cond. Matt. **1**, 6959 (1989).
[9] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).
[10] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989).
[11] N. Prokof'ev, B. Svistunov, and I. S. Tupitsyn, Phys. Lett. A **238**, 253 (1998).
[12] U. Wolff, Nucl. Phys. B **832**, 520 (2010), arXiv:1001.2231 [hep-lat].
[13] Z. Fodor, S. D. Katz, and C. Schmidt, JHEP **03**, 121 (2007), arXiv:hep-lat/0701022.
[14] S. Borsanyi and D. Sexty, Phys. Lett. B **815**, 136148 (2021), arXiv:2101.03383 [hep-lat].
[15] K. Langfeld and B. Lucini, Phys. Rev. D **90**, 094502 (2014), arXiv:1404.7187 [hep-lat].
[16] K. Langfeld, B. Lucini, and A. Rago, Phys. Rev. Lett. **109**, 111601 (2012), arXiv:1204.3243 [hep-lat].
[17] K. Langfeld, B. Lucini, R. Pellegrini, and A. Rago, Eur. Phys. J. C **76**, 306 (2016), arXiv:1509.08391 [hep-lat].
[18] K. Langfeld, PoS **LATTICE2016**, 010 (2017), arXiv:1610.09856 [hep-lat].

[19] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
[20] F. Wang and D. P. Landau, Phys. Rev. E **64**, 056101 (2001), cond-mat/0107006.
[21] E. Ising, Zeitschrift für Physik **31**, 253 (1925).
[22] L. Onsager, Phys. Rev. **65**, 117 (1944).
[23] H. Robbins and S. Monro, Ann. Math. Stat. **22**, 400 (1951).
[24] H. J. Kushner and G. G. Yin, *Stochastic Approximation Algorithms and Applications* (Springer, 1997).
[25] J. R. Blum, Ann. Math. Stat. **25**, 382 (1954).
[26] K. L. Chung, Ann. Math. Stat. **25**, 463 (1954).
[27] V. Fabian, Ann. Math. Stat. **39**, 1327 (1968).
[28] N. Garron and K. Langfeld, Eur. Phys. J. C **76**, 569 (2016), arXiv:1605.02709 [hep-lat].
[29] N. Garron and K. Langfeld, Eur. Phys. J. C **77**, 470 (2017), arXiv:1703.04649 [hep-lat].
[30] O. Francesconi, M. Holzmann, B. Lucini, and A. Rago, Phys. Rev. D **101**, 014504 (2020).
[31] J. M. Pawlowski and J. M. Urban, (2022), arXiv:2203.01243 [hep-lat].