


Reduced critical slowing down for statistical physics simulations

Kurt Langfeld 

School of Mathematics, University of Leeds, Leeds, LS2 9JT, United Kingdom

Pavel Buividovich , P. E. L. Rakow , and James Roscoe

Department of Mathematical Sciences, University of Liverpool, Liverpool, L69 7ZX, United Kingdom



(Received 22 May 2022; accepted 23 October 2022; published 16 November 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 linear-log-relaxation (LLR) method as simulation algorithm. We first demonstrate supercritical slowing down in a phase with spontaneously broken symmetry and for the heat-bath algorithms, for which autocorrelation times grow exponentially with system size. By contrast, using the magnetization as collective coordinate, we present evidence that supercritical slowing down is absent. We still observe a polynomial increase of the autocorrelation time with volume (critical slowing down), which is, however, reduced by orders of magnitude when compared to local update techniques.

DOI: [10.1103/PhysRevE.106.054139](https://doi.org/10.1103/PhysRevE.106.054139)

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 in a new configuration of degrees of freedom on the lattice. The simulations generate 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 (and assuming that the variance of A exists).

In practical Monte Carlo simulations, configurations are correlated over a characteristic number of Monte Carlo updates $t \approx \tau$, which is called autocorrelation time (we will give a proper definition below). An immediate impact is that the statistical error now scales like $\sqrt{\tau/N}$. Large autocorrelation 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 autocorrelations.

The autocorrelation 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 $\xi \rightarrow \infty$ is of crucial importance to extract physics relevant information from those computer simulations. Generating independent Markov ensembles in the case that degrees of freedom are correlated over many sites is a challenge for any algorithm and in particular, for the important class of *local update algorithms*. This challenge is reflected by the 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:

$$\tau(L) \propto L^z, \quad (\text{critical slowing down})$$

$$\tau(L) \propto e^{mL}, \quad (\text{supercritical slowing down}).$$

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

Over many decades, research has been analyzing the combination of theories and algorithms studying autocorrelations times for particular observables. For Markov chain simulation

that satisfies detailed balance, large autocorrelation 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 characterization 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 decorrelation typically correspond to long-wavelength modes of physical fields. For a free scalar field theory, a combination of particular order of updating the fields and tuning of stochastic over-relation *can* significantly reduce critical slowing down [8]. Albeit this is per se an interesting finding, we here do not consider algorithms that need significant fine tuning for reducing autocorrelations.

To alleviate the “slow mode relaxation” issue, multigrid methods have been proposed already in the late 1980s [9]. For specific models, targeted solutions can be found that either eliminate critical slowing down or strongly reduce it. Those attempts are based on a reformulation, and simulations include nonlocal updates. 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 [10] for two dimensions. *Cluster update algorithms* [11,12] generically possess a small dynamical critical exponent z and thus provide a practical solution to the critical slowing down issue. Whenever a model allows a cluster reformulation, the performance cluster algorithms are hardly outperformed by any other approach and hence are the preferred simulation method.

Lattice theories that show *spontaneous symmetry breaking* in the infinite volume limit are particularly prone to *supercritical slowing down* when simulated 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 $M(\phi) \neq 0$ [13], and $\langle M \rangle$ vanishes upon averaging over these relevant domains. Local update algorithms usually fail to induce transitions between these domains, leading to supercritical slowing down. Yang-Mills theories with a gauge group $SU(N \geq 3)$ fall into this important class of models [14]. Gauge symmetry prevents the definition of meaningful (gauge invariant) clusters and corresponding nonlocal update algorithms. We are hence turning to other, more conventional simulation techniques.

A promising class of such algorithms are multicanonical algorithms [15] and Wang-Landau techniques [16,17]. Although the algorithmic differences and similarities between both methods have been studied in the literature (see, e.g., [18]), both employ reweighing techniques with respect to a marginal distribution, which is at the heart of solving the issue of supercritical slowing down. This has been first demonstrated by Torrie and Valleau [19] in a thermodynamics setting and later by Berg *et al.* for the Ising model in [20].

At the root of supercritical slowing down is the double-peak marginal distribution $P(M)$ of the order parameter, say the magnetization M . Rather than leave it to importance sam-

pling to transition between the two equally important phases, we calculate the partition function by integrating *explicitly* over the order parameter M and stochastically over the remainder of the configuration space. To this aim, we exploit the identity

$$Z(\beta) = \int dm \rho(m),$$

$$\rho(m) = \int \mathcal{D}\phi \delta(m - M(\phi)) \exp\{\beta S(\phi)\},$$

where δ is the Dirac δ 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 [21], a recent study of the topological density in pure Yang-Mills theories [22], and the first proof of concept of solving a strong sign problem using the Z_3 theory [23].

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 this paper we explore the *linear-log-relaxation* (LLR) method [24–26], which belongs to the class of the Wang-Landau techniques. The LLR method is based upon a systematic expansion of the marginal distribution $\rho(m)$ in a given m interval, leading to stochastic nonlinear equations for the expansion parameters (see below for details). In its lowest order, the LLR approach has similarities with the “multi-magnetic ensemble” method by Berg *et al.* [20]. The LLR approach is also markedly different: it confines the MC simulation part to a window of size 2δ around a given value of the magnetization m_0 , which is a nonlocal constraint. We will be interested in the limit $\delta \rightarrow 0$.

In this paper we offer a systematic and large-scale study of the phenomenon of critical slowing down using the LLR method. Since we are interested in simulation methods, which can be applied universally to a wide range of lattice models, we benchmark our findings against those from a heat-bath approach rather than a cluster algorithm, which would be the method of choice anyway if applicable. We find evidence that *supercritical slowing down* is absent (in line with the findings from a multicanonical simulation [20]). We still find a correlation length that increases polynomially with the volume. We observe, however, that those correlations are strongly suppressed even at criticality.

II. UNDERSTANDING CRITICAL SLOWING DOWN

A. Accessing autocorrelations

The well-studied Ising model in a finite volume also serves here to illustrate the breakdown of *importance sampling* due to a failure of sampling the configuration space within an acceptable amount of computational resources. The purpose of this section is to quantify this breakdown for the popular Markov-chain Monte-Carlo 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 with periodic boundary conditions [27] throughout the paper. The partition function Z and action S are given by

$$Z = \sum_{\{s_x\}} \exp\{\beta S\}, S = \sum_{\langle xy \rangle} s_x s_y, \quad (1)$$

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

(1) Choose a site x of the lattice at random and calculate the sum over the neighboring 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 and 2 above V times 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 by the Monte Carlo time $k = 1 \dots N$. Define a sequence of random numbers for an observable $f(\{s_x\})$ by

$$f_1 \rightarrow f_2 \rightarrow \dots \rightarrow f_N, f_i = f(\{s_x\}_i).$$

(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 *magnetization per spin*,

$$\langle m \rangle = \left\langle \frac{1}{V} \sum_x s_x \right\rangle = \langle s_x \rangle,$$

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

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

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

By the law of large numbers, we find

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

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

$$\langle m \rangle \approx [m] \quad (3)$$

valid?

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

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

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

$$\epsilon^2 = [m^2] - [m]^2 = [m^2]. \quad (4)$$

We find

$$\epsilon^2 = \left[\left(\sum_{i=1}^N \sum_{\ell=1}^N m_i \right)^2 \right] = \frac{1}{N^2} \sum_{i=1}^N \sum_{\ell=1}^N [m_i m_\ell]. \quad (5)$$

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

$$[m_i m_\ell] = m_0^2 \exp\left\{-\frac{k}{\tau}\right\}, \quad k = |i - \ell| \quad (6)$$

$$m_0^2 := [m_i^2],$$

where τ is called *autocorrelation* time. This is expected to be the case for large separations k . A rather stark assumption is that the exponential behavior dominates the double sum in (5). This assumption only can be justified afterwards in the numerical experiment, but it seems to be the case for the parameter range explored in this paper. Inserting (6) into (5), the double sum can be performed analytically:

$$\epsilon^2 = \frac{m_0^2}{N^2} \sum_{\ell=1}^N \sum_{i=1}^N a^{|i-\ell|} = \frac{m_0^2}{N} \frac{1+a}{1-a} - \frac{2a m_0^2}{N^2(1-a)^2} (1-a^N), \quad (7)$$

$$a = \exp\{-1/\tau\}. \quad (8)$$

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

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

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

In 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:

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

In this case, the error is of order 1, and we cannot expect that (3) yields a meaningful approximation. Note, however, that

Eq. (10) still can provide information on the (large) autocorrelation time by virtue of the correction to the leading term even if $N \sim \tau$.

B. Symmetry breaking and ergodicity

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

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

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 sequences

$$m_1 \rightarrow m_2 \rightarrow \dots m_N \text{ and } -m_1 \rightarrow -m_2 \rightarrow \dots -m_N$$

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

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

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 $V \rightarrow \infty$. 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 realized. The critical value β_c can be calculated analytically [28], and one finds

$$\beta_c = \frac{1}{2} \ln(1 + \sqrt{2}) \approx 0.440686\dots \quad (13)$$

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 important sampling algorithms (if not all) for large enough $\beta \gg \beta_c$ 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 $N \rightarrow \infty$, 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 magnetization 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 the system

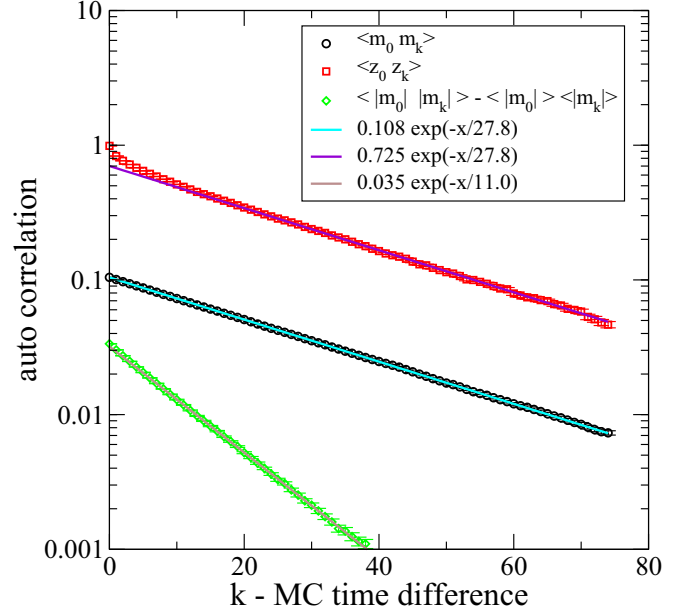


FIG. 1. Autocorrelation functions for a 12×12 Ising model at $\beta = 0.35$ as a function of the MC time difference k (see 6).

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 realize that regions of positive (negative) m_i cluster for some time. This indicates that we observe a significant autocorrelation time τ even at this small β . In order to quantify this, we present estimators for the autocorrelation 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 nonzero 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 within the symmetric phase with a moderate autocorrelation time. The simulation starts with a random spin configuration (hot-start) and initially discards 1000 configurations for thermalization. The result for the autocorrelation functions is shown in Fig. 1, right panel. Our findings suggest that the autocorrelation functions of m and z are proportional (at least for sufficiently large a MC time difference), i.e.,

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

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

We have systematically studied the error ϵ [as given by Eq. (4)] for an $L = 12$ 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 precise] to the numerical data. This yields an estimate for m_0^2 and the desirable autocorrelation time τ . Our findings are summarized in Fig. 2, left panel. For beta 0.3 and 0.35 the observed autocorrelation time is small enough so that we can observe the characteristic $1/\sqrt{N}$ behavior at large N . Note, however, that close to $\beta \approx \beta_c$, we

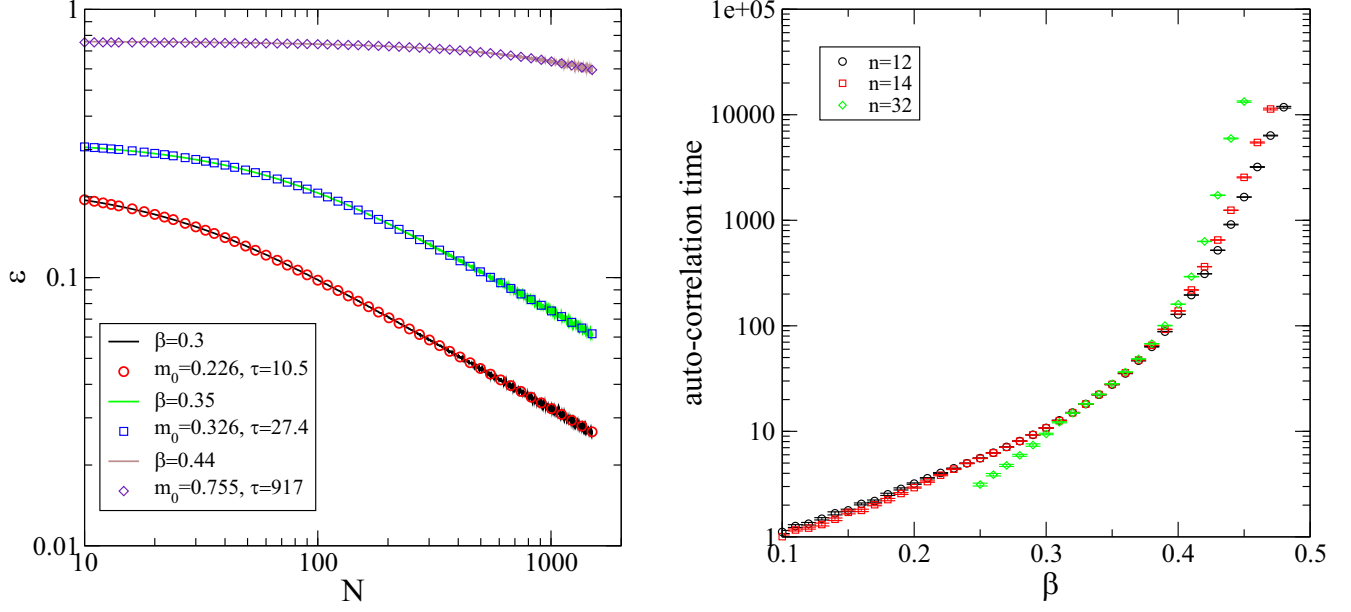


FIG. 2. Left panel: Solid lines are estimates [see (17)] for the statistical error ϵ as a function of the length N of the MCMC time series; 12×12 Ising model. Open symbols are the theoretical prediction (7). Right panel: extracted autocorrelation time as a function of β for several lattice sizes; $n \times n$ Ising model.

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 Fig. 2, right panel, shows the autocorrelation time as a function of β for the three lattice sizes 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 $\beta > \beta_c$, which corresponds to the symmetry broken phase in the infinite volume limit.

Equation (14) suggests that tunneling between the Z_2 sector is suppressed and that this suppression is at the heart of the practical ergodicity issue. For each step 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 $z_i, i = 1 \dots k + 1$, 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 autocorrelation function $\langle z_1 z_{k+1} \rangle$ is $(-1)^\ell$. Hence we find

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

Using the latter result in (14) and exploiting the connection to the autocorrelation time in (6), we find the connection between the autocorrelation time τ and sector tunneling prob-

ability p :

$$p = \frac{1}{2}(1 - e^{-1/\tau}) \approx \frac{1}{2\tau}. \quad (16)$$

The latter approximation holds for $\tau \gg 1$. For the example of the previous section, 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?

For the heat-bath algorithm, we already worked out a connection between ϵ^2 and N [see Eq. (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 magnetizations m_k , which we divide into subsequences of length N . For each subsequence, we calculate the average magnetization

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

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

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

Our numerical findings for $N = 10 \dots 1500$ appear in Fig. 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 predic-

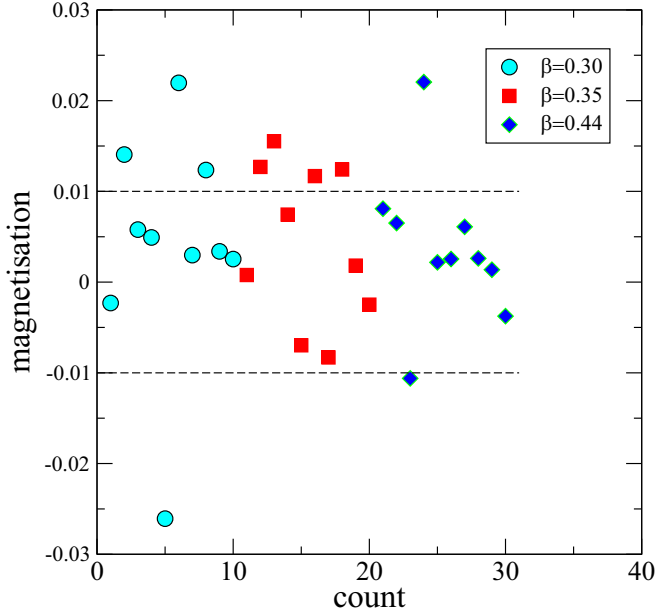


FIG. 3. Average magnetization from a MCMC time series of length N for three β [see (18) for the β - N pairs]; 12×12 Ising model.

tion (7) with respect to only two fit parameters: m_0 and τ . The agreement is excellent.

We can now ask the following question: At 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 so good that we can extrapolate to N values bigger than 1500. We find that for our lattice size $L = 12$, N has to be at least the following:

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

Note that the above N values are vastly outside the fitting range of $N = 10 \dots 1500$, 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 magnetization. We have repeated this ten 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 Fig. 3. We observed the expected behavior 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 autocorrelation time. We have done this for a range of β values and show the result in Fig. 2, right panel. We observe that the autocorrelation 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 reasonably 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 a given 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 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 center sectors is exponentially suppressed, and a changing a Z_2 sector needs resources with exponentially increasing volume. In this section we 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 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 with

$$t > 200, \quad t < t_{\max},$$

where

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

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 Fig. 4, top panel. The parameters were $L = 16, 32$ 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 are obtained by bootstrap. For the fits shown in Fig. 4, we obtained specifically

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

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 are shown in the same Fig. 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

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

Our findings are summarized in Table I. 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 a dominating exponential growth. The fits are also shown in the bottom panel of Fig. 4. They well describe the data. In particular, we thus find evidence that b_2 starts growing to nonzero values around the critical values $\beta \approx \beta_c$ for the phase transition. In the symmetric phase at $\beta = 0.43$, we find

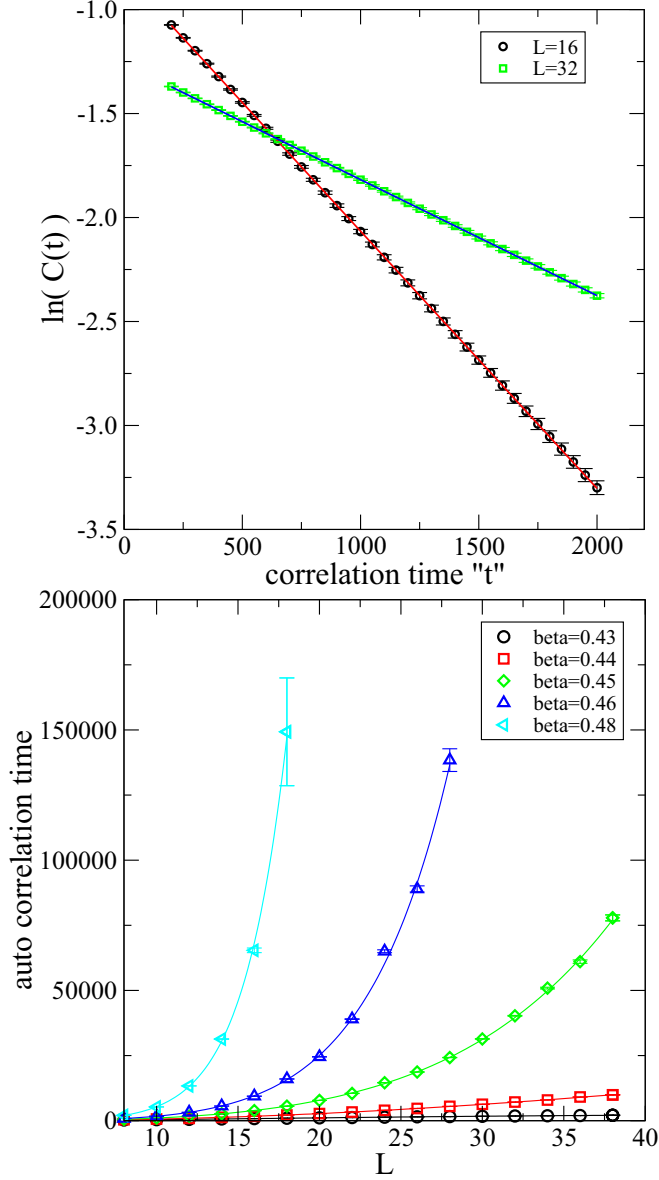


FIG. 4. Autocorrelation function as a function of Monte Carlo time t for two lattice sizes at $\beta = 0.43$ (top). The autocorrelation time for the magnetization is observable as a function of system size L for several values of β (bottom).

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

	$\ln(b_0)$	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)

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 magnetization 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 magnetization:

$$\rho(M) = \frac{1}{Z} \sum_{\{s_x\}} \delta\left(M, \sum_x s_x\right) \exp\{\beta S\}, \quad (20)$$

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

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

The magnetization is then given by

$$\langle m \rangle = \frac{\sum_M M \rho(M)}{\sum_M \rho(M)},$$

$$M = -V, -V + 2, \dots, V - 2, V. \quad (21)$$

With the normalization

$$\sum_M \rho(M) = 1, \quad (22)$$

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

$$\langle\langle f \rangle\rangle(a) = \frac{1}{\mathcal{N}} \sum_{\{s\}} f(s) e^{\beta S + a m(s)} W_\delta(m_0, m(s)),$$

$$m(s) = \sum_x s_x, \quad (23)$$

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

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

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

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

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

2. Use

$$\frac{d}{dm_0} \ln \rho(m_0) = \lim_{\delta \rightarrow 0} a(\delta, m_0) \quad (26)$$

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

3. Determine the multiplicative factor by normalizing ρ [see (22)].

The last step might be optional, since a normalization constant drops out of expectation values such as that of Eq. (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 first by Robbins and Monro [29] and then taken up by a number of authors (see [30] for a review). If $F(a)$ is a noisy estimator for

$$f(a) := \langle (m(s) - m_0) \rangle (a), \quad (27)$$

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

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

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

$$\sum_{n=1}^{\infty} \alpha_n \rightarrow \infty, \quad \sum_{n=1}^{\infty} \alpha_n^2 \rightarrow \text{finite}.$$

The sequence converges with probability 1 to the solution $a^* := a_\infty$ [31]. 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 [32] and Fabian [33] 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., Eq. (25), we can, however, find a good value κ . For small enough δ , the marginal value for the magnetization m in the window $[m_0 - \delta, m_0 + \delta]$ is Poisson distributed, i.e., $\propto \exp\{-a^* m\}$. Together with the “reweighting” 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 freedom 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:

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

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 Fig. 5. We performed 1000 independent Robbins-Monro runs to estimate the error for ϵ_a . We find optimal convergence behavior 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” needed to calculate the magnetization $\langle m \rangle$ 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 simulation parameters are “ballpark” 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 $\delta \rightarrow 0$ 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., $(m_0)_k = -32^2 + k \times 32$, $k = 1 \dots 63$.

(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 us to measure resources independent of hardware employed for the calculations. All algorithms studied here—heat-bath update, cluster algorithms, LLR method—use lattice sweeps at low levels 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

$$4 \times 20 \times 10\,000 \times 63 \, ls = 5.04 \times 10^7 \, ls. \quad (30)$$

From this data set, we can already estimate expectation values of functions of the magnetization, 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

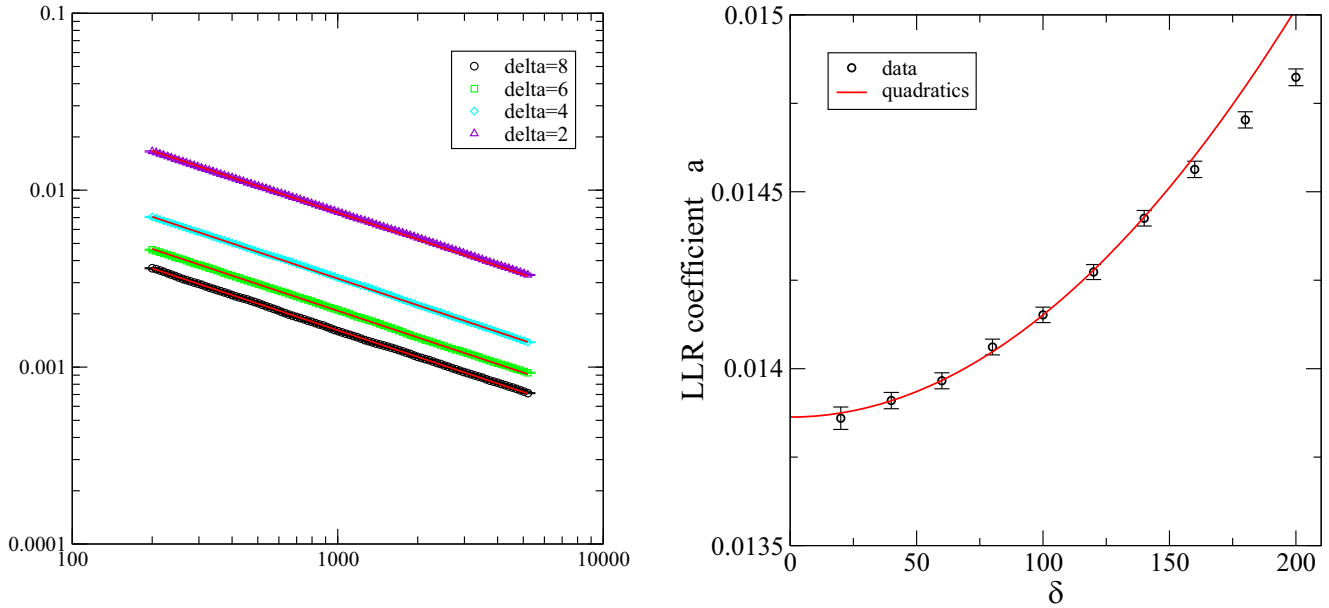


FIG. 5. Left: The error in the LLR coefficient a as a function of the number of Robbins-Monro iterations n (29). The fits correspond to a $1/\sqrt{n}$ power law. 12×12 Ising model, $\beta = 0.30$. Right: Dependence of the LLR coefficient a on δ for a 64×64 lattice near criticality ($\beta = 0.44$).

aim, we will repeat the calculation 80 times. Thus the analysis uses the resources of $5.04 \times 10^7 \times 80 \text{ ls} = 4.032 \times 10^9 \text{ ls}$, which must not be confused with the resource (30) needed to produce one sample result.

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

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

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

$$\langle m \rangle = \int m \rho(m) dm / \int \rho(m) dm. \quad (32)$$

Early studies [23–25] used a trapezium rule and summation, which leads to an accumulation of error for increasing m . Representing the function $a(m)$ by a high-degree polynomial and performing the integrations (semi-) analytically has proven very successful [26,34–36]. One can prove that the density of states for the 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 a polynomial of odd powers of m . This would lead to the exact result $\langle m \rangle = 0$.

The prime objectives here are 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 a polynomial containing even and odd powers of m . We find that a polynomial of degree of 16 represents the numerical data for a very well.

The result for $\rho(m)$ (on a logarithmic scale) is shown in Fig. 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 Fig. 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 magnetization. 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).

Here we are not primarily interested in the density of states ρ but the expectation value of the magnetization:

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

In this case, we replace step 5a by

(5b) For the sample $\rho(m)$, calculate the two integrals in (31) analytically 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 magnetization $m = M/L^2$. For the finite volume $L = 32$, we see that the most likely magnetizations 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.

Having calculated the density of states, we estimated the magnetization m using (32). The precision with which the

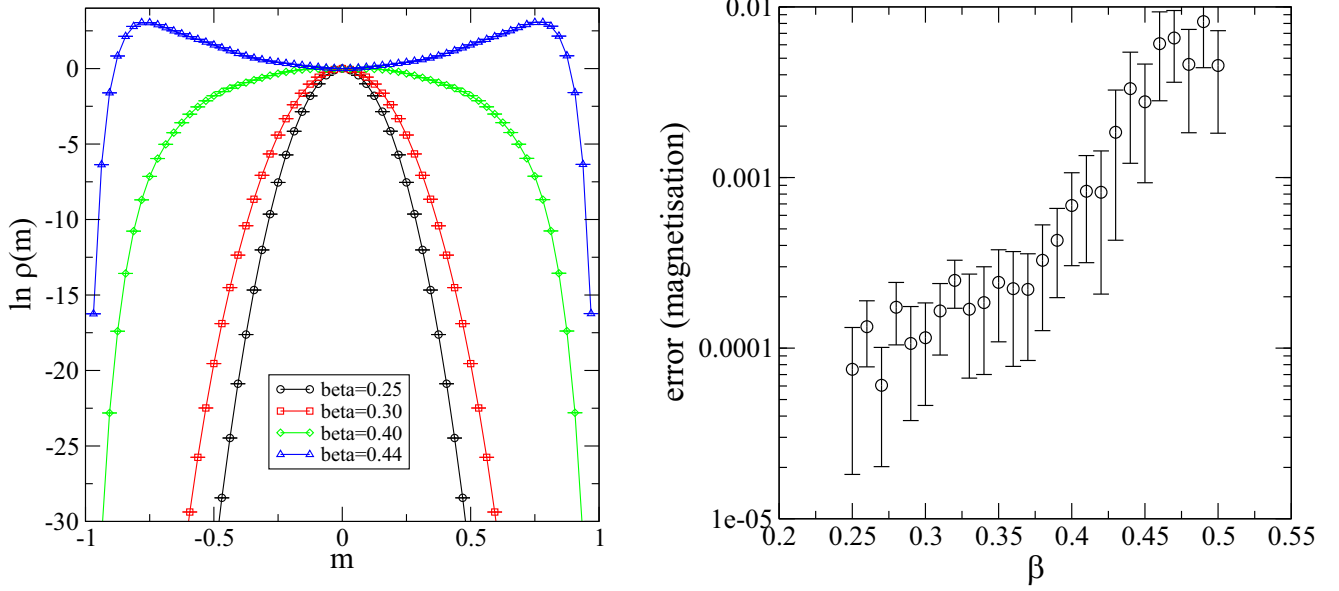


FIG. 6. Left: Log of the density of states $\rho(m)$ as a function of the (intensive) magnetization m for four β values; 32×32 Ising model. Right: Error of the magnetization (32) as a function of β .

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 Fig. 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 to 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 Eq. (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 at close to criticality “most of the time.”

We first note that the double-expectation values depend on a number of parameters that 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 $s_x \rightarrow -s_x$. Secondly, the window function $W(m_0, m(s))$ (24) is part of the probabilistic measure. It restricts spin configurations to values of the magnetization $m(s)$ 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., Eq. (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 center of the window function:

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

Our findings are summarized in Fig. 7, left panel. Indeed, we observe that those autocorrelations are highest close to $m_0 = 0$, where the system can have critical behavior.

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

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

For $p_x = 0, p_y = 0$, this quantity becomes the magnetization, 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 or decorrelation rate in local, translationally invariant quantum field theories. We thus study the autocorrelation time for the observable:

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

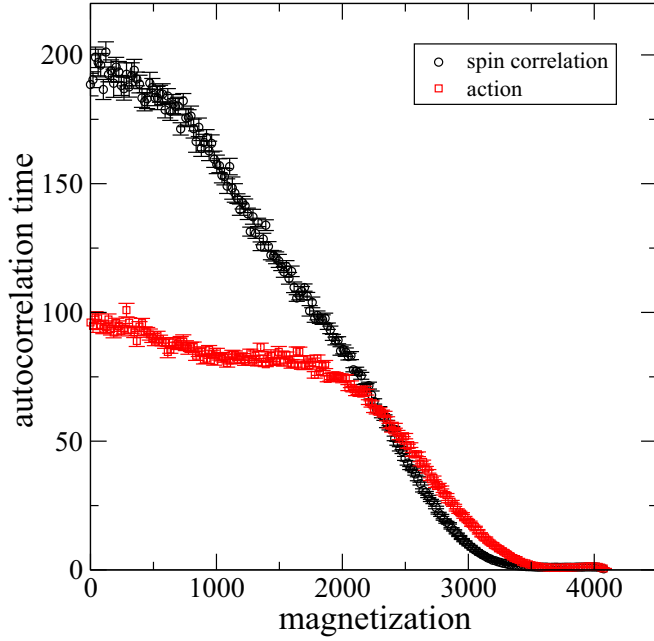


FIG. 7. Autocorrelation time for the LLR double-expectation value for the action and the spin-spin correlation as a function of m_0 . 64×64 Ising model, $\beta = 0.44$, $\delta = 40$.

To this end, we first estimate the autocorrelation function $C(t)$ of M_1 and extract the autocorrelation time by analyzing 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 holds 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 correlation 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 Fig. 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 for those of the heat-bath simulation and the magnetization M .

The central question is whether or not these autocorrelation 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 heat-bath result. Of particular interest is the coefficient b_2 , which indicates supercritical slowing down for $b_2 > 0$. Our findings are summarized in Table II.

We observe a very small coefficient b_2 when compared to the heat-bath simulation, where $b_2 \approx 0.28$ at $\beta = 0.48$. The quality is less convincing, especially for $\beta = 0.5$. Figure 9 shows two fits: the exp-power-law fit (19) and a power-law

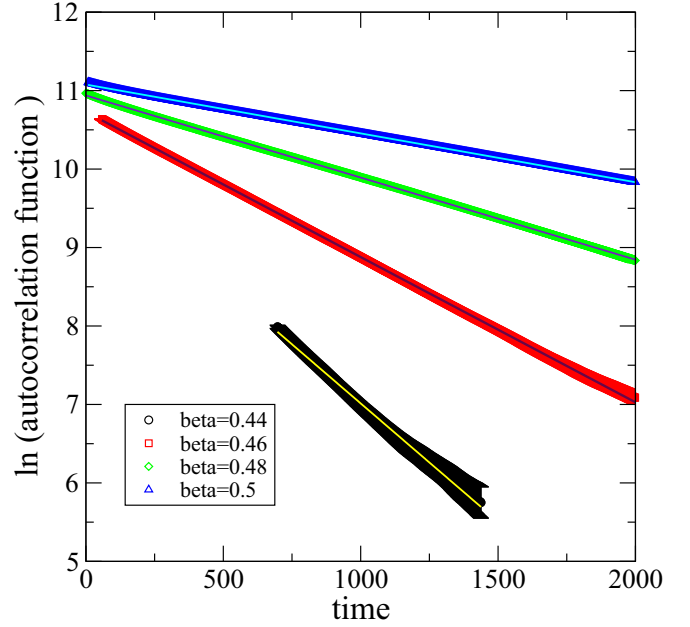


FIG. 8. Autocorrelation function for the observable M_1 (34) as a function of the autocorrelation time for four values of β . 32×32 Ising model, $m_0 = 0$.

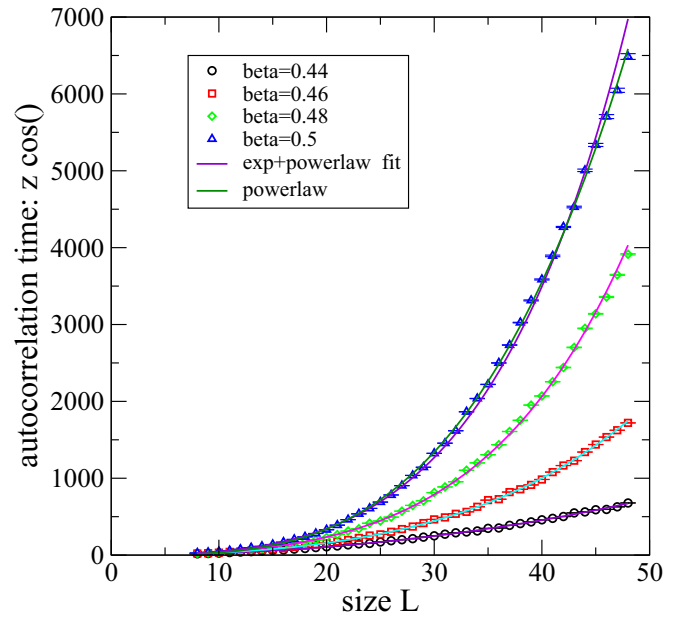


FIG. 9. Autocorrelation time for the observable M_1 (34) as a function of the system size L for four values of β and for the worst case scenario $m_0 = 0$.

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)

	$\ln(b_0)$	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)

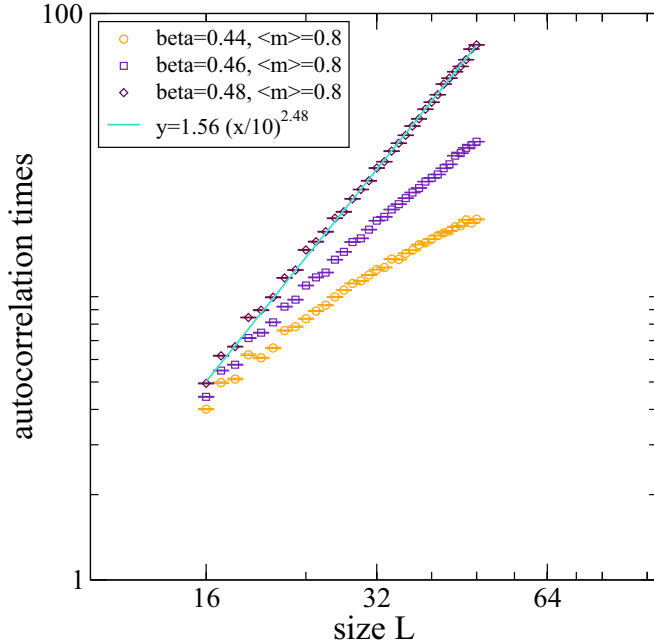


FIG. 10. A comparison of the dependence of autocorrelation time for the observable M_1 on lattice size L for the LLR approach for $m_0 = 0.8$ for several values of β .

fit $b_2 = 0$. Both fits reasonably well present the data. We are carefully optimistic that any exponential growth is a quite small rate, implying that autocorrelation 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 scales with the lattice size L . In the broken phase, say for $\beta > 0.45$, the marginal distribution for the magnetizations 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 magnetization, the dominant contributions from the LLR integration over the magnetization rises 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 Fig. 10 in the double-log scale in comparison with the $m_0 = 0$ data. We observe that autocorrelation times are orders-of-magnitude 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 subpolynomial. A 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 subpolynomial 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 magnetization *are not* affected by supercritical slowing down.

IV. DISCUSSION AND CONCLUSIONS

Local update algorithms for Markov chains of a given sample size tend to fail in exploring the full configuration space, and hence ergodicity, for theories in the regime of a spontaneously broken symmetry. In this regime, the marginal distribution of the order parameter exhibits several regions of equal stochastic importance, but importance sampling generically selects only one of these regions and fails to transition between. Consequently, the autocorrelation function rises exponentially with the system size (*supercritical slowing down*). A second question which arises is whether the autocorrelation length still rises polynomially, say at criticality (*critical slowing down*). We addressed both issues in this study.

Our approach is to decompose the configuration space into the order parameter as a collective coordinate and the hyperspace orthogonal to this mode. Wang-Landau techniques (and the LLR method, in particular) are ideally placed to integrate the slow mode explicitly, while the integration over the hyperspace is done stochastically using MCMC techniques.

In this paper we used a simple two-dimensional Ising model to demonstrate how to explore the performance of the LLR method. For the Ising model, there are efficient model-specific cluster algorithms that not only eliminate *supercritical slowing down* but also largely alleviate *critical slowing down*, as witnessed by a small dynamical critical exponent. Note, however, that cluster algorithms are only available for very specific models. The present research targets algorithms that work for a large class of models “out of the box,” without major fine-tuning.

For the Ising model, the mode that exhibits the longest autocorrelation time is the global magnetization, 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. To confirm this, we also studied the autocorrelation time for the mode with lowest nonzero momentum $p = \frac{2\pi}{L}$, where L is the linear system size. 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 the longest autocorrelation time in lattice QCD.

We found that the LLR algorithm has a potential for solving the issue of *supercritical 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. We still see a polynomial rise of the autocorrelation time with the volume at criticality (and hence *critical slowing down*), but we find that at a quantitative level the autocorrelation time is reduced by orders of magnitude when compared with that of a heat-bath simulation with the same system size (see Fig. 11).

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

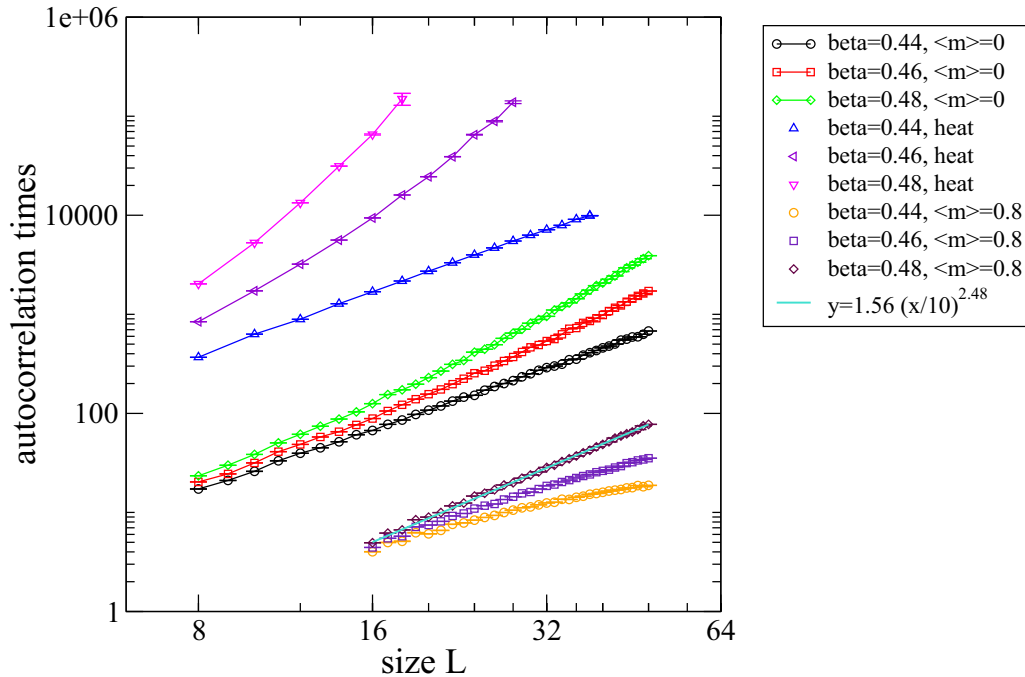


FIG. 11. A comparison of the dependence of autocorrelation time on lattice size L for the conventional heat-bath algorithm, where total magnetization has the longest autocorrelation time, and for the LLR algorithm with magnetization in the vicinity of $m_0 = 0$ and $m_0 = 0.8$, where the Fourier component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

could result in further reduction of computational time. It is also worth exploring whether the application of the LLR method to fermionic systems could reduce ergodicity issues related to zeros of the fermionic determinant. Finally, in a recent paper [37] it was suggested that normalizing flows can eliminate the need to integrate the density of states over m altogether, thus yielding an even larger speedup for Monte Carlo simulations. It would be interesting to see to what ex-

tent normalizing flows can further reduce the critical slowing down in our situation.

ACKNOWLEDGMENT

The numerical simulations were undertaken on ARC4, part of the High Performance Computing facilities at the University of Leeds, UK.

-
- [1] H. J. Rothe, *Lattice Gauge Theories*, 4th ed. (World Scientific, Singapore, 2012).
 - [2] K. Binder and D. Heermann, *Monte Carlo Simulation in Statistical Physics*, 4th ed. (Springer, Berlin, 2002), pp. XII, 180 S.
 - [3] F. M. Dekking, *A Modern Introduction to Probability and Statistics Understanding Why and How* (Springer, London, 2005).
 - [4] S. Schaefer, R. Sommer, and F. Virota (ALPHA), *Nucl. Phys. B* **845**, 93 (2011).
 - [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).
 - [7] R. Brower, S. Chandrasekharan, J. W. Negele, and U. J. Wiese, *Phys. Lett. B* **560**, 64 (2003).
 - [8] I. Horváth and A. Kennedy, *Nucl. Phys. B* **510**, 367 (1998).
 - [9] E. P. Stoll, *J. Phys.: Condens. Matter* **1**, 6959 (1989).
 - [10] U. Wolff, *Nucl. Phys. B* **832**, 520 (2010).
 - [11] R. H. Swendsen and J.-S. Wang, *Phys. Rev. Lett.* **58**, 86 (1987).
 - [12] U. Wolff, *Phys. Rev. Lett.* **62**, 361 (1989).
 - [13] K. Binder, *Z. Phys. B* **43**, 119 (1981).
 - [14] K. Holland and U.-J. Wiese, The center symmetry and its spontaneous breakdown at high temperatures, in *At The Frontier of Particle Physics* (World Scientific, 2001), pp. 1909–1944.
 - [15] A. Billoire, T. Neuhaus, and B. Berg, *Nucl. Phys. B* **396**, 779 (1993).
 - [16] F. Wang and D. P. Landau, *Phys. Rev. Lett.* **86**, 2050 (2001).
 - [17] F. Wang and D. P. Landau, *Phys. Rev. E* **64**, 056101 (2001).
 - [18] C. Junghans, D. Perez, and T. Vogel, *J. Chem. Theory Comput.* **10**, 1843 (2014).
 - [19] G. Torrie and J. Valleau, *J. Comput. Phys.* **23**, 187 (1977).
 - [20] B. A. Berg, U. Hansmann, and T. Neuhaus, *Phys. Rev. B* **47**, 497 (1993).
 - [21] Z. Fodor, S. D. Katz, and C. Schmidt, *J. High Energy Phys.* **03** (2007) 121.
 - [22] S. Borsanyi and D. Sexty, *Phys. Lett. B* **815**, 136148 (2021).
 - [23] K. Langfeld and B. Lucini, *Phys. Rev. D* **90**, 094502 (2014).
 - [24] K. Langfeld, B. Lucini, and A. Rago, *Phys. Rev. Lett.* **109**, 111601 (2012).
 - [25] K. Langfeld, B. Lucini, R. Pellegrini, and A. Rago, *Eur. Phys. J. C* **76**, 306 (2016).

- [26] K. Langfeld, *PoS (LATTICE2016)*, 10 (2017).
- [27] E. Ising, *Z. Phys.* **31**, 253 (1925).
- [28] L. Onsager, *Phys. Rev.* **65**, 117 (1944).
- [29] H. Robbins and S. Monro, *Ann. Math. Stat.* **22**, 400 (1951).
- [30] H. J. Kushner and G. G. Yin, *Stochastic Approximation Algorithms and Applications* (Springer, New York, 1997).
- [31] J. R. Blum, *Ann. Math. Stat.* **25**, 382 (1954).
- [32] K. L. Chung, *Ann. Math. Stat.* **25**, 463 (1954).
- [33] V. Fabian, *Ann. Math. Stat.* **39**, 1327 (1968).
- [34] N. Garron and K. Langfeld, *Eur. Phys. J. C* **76**, 569 (2016).
- [35] N. Garron and K. Langfeld, *Eur. Phys. J. C* **77**, 470 (2017).
- [36] O. Francesconi, M. Holzmann, B. Lucini, and A. Rago, *Phys. Rev. D* **101**, 014504 (2020).
- [37] J. M. Pawłowski and J. M. Urban, *arXiv:2203.01243*.