Testing the event-chain algorithm in asymptotically free models

Martin Hasenbusch

Institut für Physik, Humboldt-Universität zu Berlin, Newtonstr. 15, D-12489 Berlin, Germany

Stefan Schaefer

John von Neumann Institute for Computing (NIC), DESY Platanenallee 6, D-15738 Zeuthen, Germany

(Received 6 July 2018; published 4 September 2018)

We apply the event-chain algorithm proposed by Bernard *et al.* in 2009 to toy models of lattice QCD. We give a formal proof of stability of the algorithm. We study its performance at the example of the massive Gaussian model on the square and the simple cubic lattice, the $O(3)$-invariant nonlinear σ-model on the square lattice, and the $SU(3) \times SU(3)$ principal chiral model on the square lattice. In all these cases we find that critical slowing down is essentially eliminated.

DOI: 10.1103/PhysRevD.98.054502

I. INTRODUCTION

Monte Carlo simulation of statistical field theory systems is a standard tool in their study. While the development of algorithms has a long tradition, it is still an active field and source of substantial progress. Here we test a new class of recently proposed algorithms, the so-called event-chain Monte Carlo [1–4], in systems with asymptotic freedom. In particular, we study the two-dimensional $O(N)$-invariant nonlinear σ-model and the two-dimensional $SU(N) \times SU(N)$ principal chiral model. In our simulations, we take $N = 3$ in both cases. As a preliminary step, we study free field theory in two and three dimensions.

The cost of a simulation is determined essentially by two factors: the numerical cost of one update of the whole system and the autocorrelation time τ in units of these updates. For definitions of the autocorrelation time see Appendix B. It typically scales with a power of the correlation length ξ,

$$\tau \propto \xi^z. \tag{1}$$

The dynamical critical exponent z therefore characterizes the algorithm's performance close to criticality. Algorithms, such as the local Metropolis algorithm, which essentially perform a random walk in configuration space have $z \approx 2$, whereas for some models also update strategies whose autocorrelation times increase much slower with the correlation length have been devised. Examples are the cluster [5,6] and multigrid [7] algorithms. They are characterized by a coherent update of a large fraction of the field variables in one step. Such advanced algorithms, however, rely on special features of the theory to be simulated and are difficult to generalize: Algorithm performance depends strongly on the model under investigation. For a new proposal, it is therefore pivotal to test it in many cases to understand better its dynamics.

An interesting new entry in the toolbox is the so-called event-chain algorithm pioneered in Refs. [1–3]. It is basically a walker on the lattice which updates the local field variable in Monte Carlo time until a so-called event occurs, with probabilities given by the theory. Then it moves to a neighboring site, which is determined by the event. The details are given below. It is remarkable that for some theories this leads to a small value of z and not to the random walk one might naively expect.

II. EVENT DRIVEN ALGORITHM

In this section, the algorithm is discussed for a model with a single real variable at each site of the lattice. In order to apply the algorithm to more complicated models, the simple model is embedded, as we explain below. Note that the idea of embedding is also used in the case of cluster [6,8] and multigrid [9] algorithms. To define the algorithm, one first discusses infinitesimal updates with a step size $\epsilon > 0$. These updates are integrated analytically, leading to the version of the algorithm that is implemented at the end.

We consider a general scalar field theory with the action

$$S([\phi]) = \sum_{x\in\Lambda} s_x(\phi_x) + \sum_{\langle x,y\rangle} s_{\langle x,y\rangle}(\phi_x, \phi_y), \tag{2}$$

where $s_x(\phi_x)$ and $s_{\langle x,y\rangle}(\phi_x, \phi_y)$ are functions of the real variables ϕ_x. Here we consider analytic functions. However the example of hard spheres considered in Ref. [1] shows that this requirement can be relaxed. The collection of all sites of the lattice is denoted by Λ and $\langle x, y\rangle$ is a pair of interacting sites. In our numerical study, we consider nearest neighbor interactions only. We require that

$$\frac{\partial s_{\langle x,y\rangle}(\phi_x, \phi_y)}{\partial \phi_x} = -\frac{\partial s_{\langle x,y\rangle}(\phi_x, \phi_y)}{\partial \phi_y} \tag{3}$$

holds for the interaction terms.

The configuration space ϕ is now extended by two "lifting" variables: the position of the walker $\tilde{x} \in \Lambda$ at which field updates are performed and the direction $\sigma \in \{-1, 1\}$ of these changes. Both are drawn from a uniform distribution, in particular, uncorrelated to the fields ϕ. Let us denote an enlarged configuration by

$$X = (\phi, \sigma, \tilde{x}). \tag{4}$$

The action pertaining to a single site, keeping the fields at all other sites fixed, can be written as

$$S_x(\phi_x) = s_x(\phi_x) + \sum_y s_{\langle x,y\rangle}(\phi_x, \phi_y) \equiv \sum_{i=0}^{n} s_{i,x}(\phi_x), \tag{5}$$

where $i = 0$ indicates the single site term and $i = 1, 2, \ldots, n$ labels the interaction partners y of x.

The enlarged configuration X is always changed. Either $\phi_{\tilde{x}}$ is replaced by $\phi_{\tilde{x}} + \sigma\epsilon$, σ is replaced by $-\sigma$, or $\tilde{x}$ is replaced by one of its interacting partners y_i. The decisions are taken for each term of the single site action (5) separately. With the standard Metropolis probability, applied to $s_{i,\tilde{x}}$

$$p_i = \min\left[1, \exp[-s_{i,\tilde{x}}(\phi_{\tilde{x}} + \sigma\epsilon) + s_{i,\tilde{x}}(\phi_{\tilde{x}})]\right] \tag{6}$$

the proposal $\phi_{\tilde{x}} + \sigma\epsilon$ is accepted. Expanding in ϵ we get

$$p_i = \begin{cases} 1 - \sigma\frac{\partial s_{i,\tilde{x}}(\phi_{\tilde{x}})}{\partial \phi_{\tilde{x}}}\epsilon + O(\epsilon^2) & \text{for } \sigma\frac{\partial s_{i,\tilde{x}}(\phi_{\tilde{x}})}{\partial \phi_{\tilde{x}}} > 0 \\ 1 + O(\epsilon^2) & \text{for } \sigma\frac{\partial s_{i,\tilde{x}}(\phi_{\tilde{x}})}{\partial \phi_{\tilde{x}}} \leq 0. \end{cases} \tag{7}$$

The update $\phi_{\tilde{x}} + \sigma\epsilon$ is rejected, if it is rejected by at least one $s_{i,\tilde{x}}$. If the proposal is not accepted the following is done: In the case of $i = 0$, σ is replaced by $-\sigma$, while for $i > 0$ the walker is set to the site y_i. At finite ϵ there is a probability $\mathrm{O}(\epsilon^2)$ that the proposal is rejected by more than one $s_{i,\tilde{x}}$. Hence for a sequence of n updates with $t = n\epsilon$ finite, the probability for conflicting decisions is $\mathrm{O}(\epsilon)$ and hence the algorithm becomes well defined in the limit $\epsilon \to 0$. A proof of the correctness of the algorithm is given in Appendix A.

The algorithm that is implemented in the program is obtained by integrating infinitesimal steps over a finite fictitious time t. The probability that there is no rejection by $s_{i,\tilde{x}}$ in $n = t/\epsilon$ subsequent steps is

$$P_i(\phi_{\tilde{x}} + \sigma t \leftarrow \phi_{\tilde{x}}) = \prod_{m=1}^{n} p_i(\phi_{\tilde{x}} + m\sigma\epsilon \leftarrow \phi_{\tilde{x}} + (m-1)\sigma\epsilon). \tag{8}$$

Taking the logarithm and the limit $\epsilon \to 0$ one arrives at

$$\ln P_i(t) = \int_0^t \mathrm{d}\psi \min[0, -\sigma s'_{i,\tilde{x}}(\phi_{\tilde{x}} + \sigma\psi)] =: -\Delta E_i(t). \tag{9}$$

Performing the integral we get

$$\Delta E_i(t) = \sigma \sum_k [s_{i,\tilde{x}}(\phi_{1,k}) - s_{i,\tilde{x}}(\phi_{0,k})], \tag{10}$$

where k labels the intervals with $\sigma s'_{i,\tilde{x}}$ being positive throughout within $[\min[\phi_{\tilde{x}}, \phi_{\tilde{x}} + \sigma t], \max[\phi_{\tilde{x}}, \phi_{\tilde{x}} + \sigma t]]$. $\phi_{0,k}$ and $\phi_{1,k}$ are the lower/upper and upper/lower bounds of these intervals, depending on the value of σ. The field $\phi_{\tilde{x}}$ is updated until the update is rejected by one of the $s_{i,\tilde{x}}$. The times $t_i^{(\mathrm{event})}$ when these events occur are determined in the following way: One draws a uniform random number $r_i \in (0, 1]$ for each i to fix $P_i(t)$. We arrive at Eq. (11) of [2],

$$\Delta E_i(t_i^{(\mathrm{event})}) = -\ln r_i, \tag{11}$$

which has to be solved for $t_i^{(\mathrm{event})}$. The $i = i_{\min}$ with the smallest $t_i^{(\mathrm{event})}$ makes the race: $t^{(\mathrm{event})} = \min_i[t_i^{(\mathrm{event})}]$. The field is updated $\phi_{\tilde{x}} \to \phi_{\tilde{x}} + \sigma t^{(\mathrm{event})}$. For $i_{\min} = 0$ we replace σ by $-\sigma$ and $\tilde{x}$ remains unchanged. Or else, for $i_{\min} > 0$ the walker assumes the new position $y_{i_{\min}}$ and σ keeps its sign.

The algorithm evolves in a fictitious time t_{MC}. For each event the fictitious time increases as $t_{MC} \to t_{MC} + t^{(\mathrm{event})}$. A sequence of updates is started at $t_{MC} = 0$ by drawing $\tilde{x}$ and σ from a uniform distribution. The sequence of updates is stopped at some fixed t_f. To this end, events are generated until t_{MC} overshoots t_f for the first time. This last event is not taken into account in the update of the field and t_{MC} is not increased by this last $t^{(\mathrm{event})}$. Instead $\phi_{\tilde{x}} \to \phi_{\tilde{x}} + \sigma(t_f - t_{MC})$ is performed. Furthermore, measurements should be performed in equal intervals of the fictitious time. Mostly, we performed a measurement after a complete update sequence of the length t_f or after a fixed number of these sequences. Note that measuring at events leads to a bias that, at least for small system sizes, can be easily seen in the averages of estimators.

In order to arrive at a formal proof of ergodicity one has to introduce some randomness in the evolution time t_f. An alternative would be to add Metropolis updates that ensure ergodicity. In our explorative study here, similar to Refs. [1–3], we simply ignore this question.

III. THE MODELS

We consider square or simple cubic lattices. The sites of the lattice are denoted by $x = (x_0, x_1, \ldots, x_{d-1})$ with $x_i \in \{0, 1, 2, \ldots L_i - 1\}$, where d is the dimension of the system, setting the lattice spacing to unity, $a = 1$. The directions of the lattice are denoted by $\mu \in \{0, 1, \ldots, d-1\}$. There are $2d$ nearest neighbors $y_i = x \pm \hat{\mu}$, where $\hat{\mu}$ is a unit vector in μ-direction. In our simulations we take $L_0 = L_1 = \cdots = L_{d-1} = L$ throughout.

A. Free field theory

First we study the two- and three-dimensional scalar field theory. It is defined by the action

$$S = \frac{1}{2}\sum_{x,\mu}(\phi_x - \phi_{x+\hat{\mu}})^2 + \frac{m^2}{2}\sum_x \phi_x^2 + \frac{\lambda}{4!}\sum_x \phi_x^4. \qquad (12)$$

In the following we only discuss the free case $\lambda = 0$.

Here the functions (9) are easy to evaluate. For the single site term of the action we get

$$\Delta E_0(t) = \frac{1}{2}m^2 \begin{cases} 0 & \text{for } \sigma\phi_{\tilde{x}} < 0, t < -\sigma\phi_{\tilde{x}} \\ (\phi_{\tilde{x}} + \sigma t)^2 & \text{for } \sigma\phi_{\tilde{x}} < 0, t \geq -\sigma\phi_{\tilde{x}} \\ (\phi_{\tilde{x}} + \sigma t)^2 - \phi_{\tilde{x}}^2 & \text{for } \sigma\phi_{\tilde{x}} \geq 0. \end{cases} \qquad (13)$$

For $i > 0$ we get

$$\Delta E_i(t) = \frac{1}{2} \begin{cases} 0 & \text{for } \sigma\Delta\phi_{\tilde{x},i} < 0, t < -\sigma\Delta\phi_{\tilde{x},i} \\ (\Delta\phi_{\tilde{x},i} + \sigma t)^2 & \text{for } \sigma\Delta\phi_{\tilde{x},i} < 0, t \geq -\sigma\Delta\phi_{\tilde{x},i} \\ (\Delta\phi_{\tilde{x},i} + \sigma t)^2 - \Delta\phi_{\tilde{x},i}^2 & \text{for } \sigma\Delta\phi_{\tilde{x},i} \geq 0, \end{cases} \qquad (14)$$

where $\Delta\phi_{\tilde{x},i} = \phi_{\tilde{x}} - \phi_{\tilde{x}+\hat{\mu}} = \phi_{\tilde{x}} - \phi_{y_i}$. There is always a solution for a positive t.

B. Nonlinear σ-model

The action of the nonlinear $O(N)$-invariant σ-model is

$$S = -\beta\sum_{x,\mu}\vec{s}_x \cdot \vec{s}_{x+\hat{\mu}}, \qquad (15)$$

where $\vec{s}_x$ is a unit vector with N real components. We study the model on the square lattice for $N = 2$ and 3. For $N = 2$ we get the so-called XY-model. It undergoes a Kosterlitz-Thouless phase transition at $\beta_{\mathrm{KT}} = 1.1199(1)$ [10]. For temperatures above the transition, there is a finite correlation length. At lower temperatures there is no ordering and the two-point correlation function is decaying with a power law. For references see e.g., [11]. For the event-chain algorithm it is useful to write the XY-model in terms of angles α_x,

$$S = -\beta\sum_{x,\mu}\cos(\alpha_x - \alpha_{x+\hat{\mu}}). \qquad (16)$$

The pair terms of the action are

$$s_{i,\tilde{x}} = -\beta\cos(\alpha_{\tilde{x}} - \alpha_{y_i}). \qquad (17)$$

In the update, $\alpha_{\tilde{x}}$ is incremented. The functions (9) are given in Ref. [2]. To simplify the notation, let us assume $\sigma = 1$ in the following. Let us define

$$\delta_i = \alpha_{\tilde{x}} - \alpha_{y_i} - 2m\pi, \qquad (18)$$

where the integer m is chosen such that $-\pi < \delta_i \leq \pi$.

$$\Delta E_i(t) = \beta \begin{cases} 2n & \text{for } \delta_i < 0, \delta_i + t - 2n\pi < 0, \\ 2n + 1 - \cos(\delta_i + t) & \text{for } \delta_i < 0, \delta_i + t - 2n\pi \geq 0, \\ 2(n-1) + \cos(\delta_i) - 1 & \text{for } \delta_i \geq 0, \delta_i + t - 2n\pi < 0, \\ 2n + \cos(\delta_i) - \cos(\delta_i + t) & \text{for } \delta_i \geq 0, \delta_i + t - 2n\pi \geq 0, \end{cases} \qquad (19)$$

where the integer $n \geq 0$ is chosen such that $-\pi < \delta_i + t - 2n\pi \leq \pi$. Note that $2n$ is the contribution from n complete cycles, where $\cos(0) - \cos(\pi) = 2$. For an illustration see Fig. 4 of Ref. [2].

For $N = 3$ the model has a finite correlation length at any finite value of β. The model is asymptotically free. The divergence of the correlation length as $\beta \to \infty$ is governed by the so-called β-function. For a discussion of the physics of this model see e.g., [12].

In the literature there are a number of Monte Carlo studies of this lattice model. It can be efficiently simulated by using the cluster algorithm [13] and the multigrid algorithm [9]. Also the microcanonical overrelaxation algorithm [14] has been applied successfully.

For simulations with the worm algorithm see Ref. [15]. The simulation for $N > 2$ with the event-chain Monte Carlo algorithm is performed, by an embedding of a generalized XY model. To this end, for one sequence of updates, one picks out a pair (l, k) of components from the N components of the field. In the update, only rotations in this plane are performed. The algorithm is run as for the XY model, with the exception that in Eq. (19) the prefactor β is replaced by $\sqrt{s^2_{l,\tilde{x}} s^2_{k,\tilde{x}}}\sqrt{s^2_{l,y_i} s^2_{k,y_i}}\beta$.

C. $SU(N) \times SU(N)$ principal chiral model

The principal chiral model on the square lattice is defined by the action

$$S = -\beta \sum_{x,\mu} \mathrm{Retr} U_x U_{x+\hat{\mu}} \tag{20}$$

where $U_x \in \mathrm{SU}\,(N)$. The action restricted to a single site is

$$S_x = -\beta \sum_{\pm\hat{\mu}} \mathrm{Retr} U_x U^\dagger_{x+\hat{\mu}} \equiv \sum_{i=1}^{2d} s_{x,i}. \tag{21}$$

The model is asymptotically free. For a discussion of the physics of this model see e.g., [16]. No efficient implementation of the cluster algorithm has been devised for this model so far. However the multigrid algorithm has been implemented successfully [16,17]. To this end, in Ref. [16], for one sequence of updates, a one-parameter subgroup of $SU(\mathrm{N})$ is considered. Here we apply the same idea. Following [16], a general one-parameter update can be composed of a random $\mathrm{SU}(N)$ rotation and a rotation by a variable angle along one, fixed element λ of the algebra. With

$$\lambda = \begin{pmatrix} i & 0 & \ldots & \\ 0 & -i & 0 & \ldots \\ 0 & \ldots & & \end{pmatrix}$$

the updates read

$$U \to e^{tR\lambda R^\dagger} U = Re^{t\lambda} R^\dagger U,$$

leading to an embedded action for site x and component i,

$$s_{i,x}(t) = -\beta[a\cos t + b\sin t] = -\beta c\cos(t - t_0)$$

with

$$a = \mathrm{Re}([W_i]_{11} + [W_i]_{22}) \quad \text{and} \quad b = \mathrm{Im}([W_i]_{11} - [W_i]_{22}), \tag{22}$$

where $W_i = R^\dagger U_x U^\dagger_{y_i} R$. Furthermore $c = \sqrt{a^2 + b^2}$ and $t_0 = \arctan(b/a)$.

IV. NUMERICAL RESULTS

Below we discuss our numerical results for the scalar free field theory, the $O(3)$-invariant nonlinear σ-model and the $SU(3) \times SU(3)$-invariant principal chiral model. We abstain from a discussion of our simulations of the two-dimensional XY-model, since our results are fully consistent with those of Ref. [18].

A. Scalar free field theory

We simulated the Gaussian model on the square and simple cubic lattice. In our simulations, we have taken $L \propto m^{-1}$, evolving the fictitious Monte Carlo time for the period t_f using the event-chain algorithm. Once this is reached, a new site $\tilde{x}$ and direction σ are selected. Note that t_f is the only free parameter of the algorithm.

Let us define the quantities that we measured in our simulations. The first two quantities are related to the terms in the action,

$$O_1 = \frac{1}{2dL^d} \sum_{\langle xy \rangle} (\phi_x - \phi_y)^2, \qquad O_2 = \frac{1}{2L^d} \phi^2_x. \tag{23}$$

Furthermore we determined the magnetization and a proxy of the second moment correlation length,

$$O_3 = \frac{1}{2}\left(\sum\nolimits_x \phi_x\right)^2, \qquad O_4 = \sum_{i=0}^{d-1} \tilde{\phi}_i \tilde{\phi}^*_i \tag{24}$$

with the Fourier transform $\tilde{\phi}_i = \sum_x \exp(-i2\pi x_i/L)\phi_x$.

In the Gaussian model, these quantities can be easily computed exactly by Fourier transformation. We compared the outcome of our simulations with these exact results. The largest deviation had been 3.4 standard deviations for a single observable.

We performed a measurement after n_m sequences of the length t_f. For a given L we kept $t_\mathrm{meas} = n_\mathrm{m} t_f$ constant. It is chosen such that the integrated autocorrelation times in units of t_meas are of $O(10)$. Studying local algorithms, such as the local Metropolis algorithm, one typically quotes the autocorrelation times in units of sweeps over the whole lattice. In our case, the artificial time that is needed on average for L^d events could be used as an analog. As discussed below, we find that the number of events per unit of the artificial time is slightly larger than one for both the square as well as the simple cubic lattice. Similar observations are made in the case of the $O(3)$-invariant nonlinear σ-model and the $SU(3) \times SU(3)$-invariant principal chiral model.

1. Two dimensions

In the two-dimensional case we simulated the linear lattice sizes $L = 32$, 64, 128, and 256. Correspondingly the masses are taken as $m = 0.3$, 0.15, 0.075, and 0.0375,

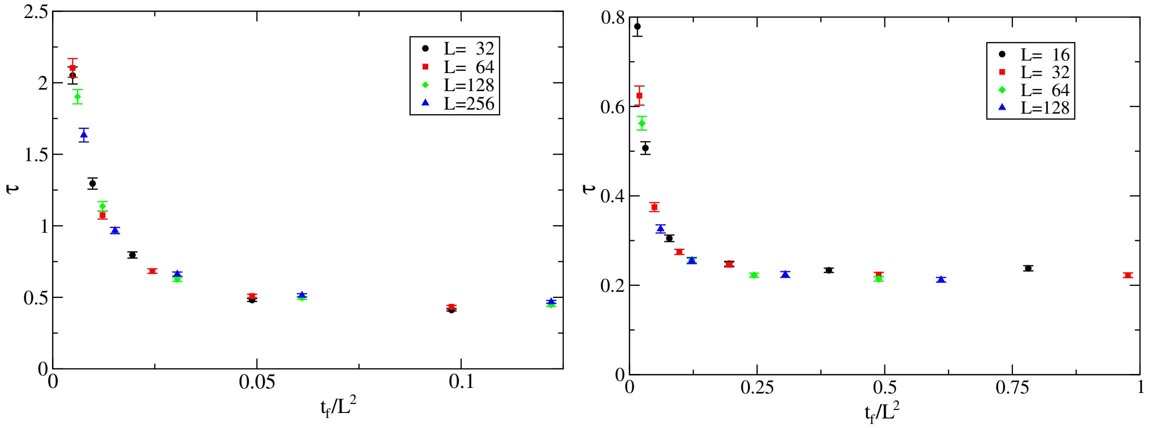

FIG. 1. Integrated autocorrelation time for the observable O_3, left the two-dimensional Gaussian model, right the three-dimensional version. The integrated autocorrelation times are given in units of L^d.

respectively. Each series of simulations consists of 10^6 measurements, separated by $t_{\mathrm{meas}} = 100$, 400, 2000 and 8000 for $L = 32$, 64, 128, and 256, respectively. We focus on the dependence of the autocorrelations on the length t_f of the update sequence.

In the left part of Fig. 1 we give the integrated autocorrelation time of O_3. We determined the integrated autocorrelation time in units of measurements. In the plot, we give τ_{int} in units of L^2. To this end, we multiply the numbers that we compute by t_{meas}/L^2. Plotting the autocorrelation time as a function of t_f/L^2, we see a reasonable collapse of the data for the four different lattice sizes. We conclude that t_f should be chosen such that $t_f \propto m^{-2} \propto L^2$. The integrated autocorrelation time of O_3 decreases with increasing t_f and seems to approach a plateau value for large t_f. We cannot exclude that τ_{int} increases for very large t_f again, as it is the case for the principal chiral model, discussed below.

A perfect collapse of the data for the different lattices sizes would also imply that critical slowing down is eliminated completely. However we see a slight increase of τ_{int} with increasing L. This is studied below for a fixed value of t_f/L^2 using higher statistics. Now let us briefly comment on the behavior of τ_{int} for the observables that we do not show in the plot. The behavior of τ_{int} for O_4 is similar to that of O_3. In the case of O_1 and O_2 the behavior is different. For O_1 we see an increase of τ_{int} with increasing t_f. This increase is not dramatic and τ_{int} seems to approach a plateau. For O_2 we see a decrease of τ_{int} with increasing t_f. It is however less pronounced as for O_3 and O_4.

Next we performed simulations for $t_f/L^2 = 125/4096 \approx 0.030517578\ldots$ fixed. Measurements were performed after $t_{\mathrm{meas}} = 2t_f$. Here we performed 5×10^6 measurements for each lattice size. In addition to $L = 32$, 64, 128, and 256 we have simulated $L = 512$. The simulation for $L = 512$ took 8 hours on one core of a Intel(R) Xeon(R) CPU E5-2660 v3 2.60 GHz. The results for the integrated autocorrelation times are summarized in Table I.

Looking at the numbers we see that for each observable, the integrated autocorrelation time is increasing with increasing lattice size. Fitting the numbers for the observable O_1 with the ansatz $\tau_{\mathrm{int},1} = cL^z$, taking into account the linear lattice sizes $L \geq 32$, we get $z = 0.112(4)$ and $\chi^2/\mathrm{d.o.f.} = 1.15$. Fits for the other three observables result in even smaller values for z. Fits with the ansatz $\tau_{\mathrm{int}} = a + c\ln(L)$ seem to be even more adequate. For example, fitting all data for $\tau_{\mathrm{int},1}$ results in $c = 1.05(2)$ and $\chi^2/\mathrm{d.o.f.} = 0.68$. We conclude that slowing down is essentially eliminated.

The CPU time that is used is proportional to the number of events. There is a small dependence on the mass and the lattice size of the average number of events within one update sequence divided by t_f. Roughly this ratio equals 1.128 for all our masses and lattice sizes.

Note that our results are consistent with those of Ref. [18], where the massless case was studied.

TABLE I. Integrated autocorrelation times for $t_f/L^2 = 125/4096 \approx 0.030517578\ldots$. The distance between two subsequent measurements is $t_{\mathrm{meas}} = 2t_f$. For each L, 5×10^6 measurements are performed. The τ_{int} are given in units of t_{meas}. To get the same normalization as in Fig. 1, the number has to be multiplied by $t_{\mathrm{meas}}/L^2 = 250/4096 \approx 0.061035156\ldots$. The index i of $\tau_{\mathrm{int},i}$ specifies the observable.

L	$\tau_{\mathrm{int},1}$	$\tau_{\mathrm{int},2}$	$\tau_{\mathrm{int},3}$	$\tau_{\mathrm{int},4}$
32	7.23(5)	4.80(3)	3.70(3)	3.56(2)
64	8.03(5)	4.80(3)	3.90(4)	3.79(3)
128	8.74(5)	4.95(3)	4.10(4)	3.99(3)
256	9.47(6)	5.09(3)	4.26(4)	4.18(3)
512	10.13(6)	5.16(3)	4.42(4)	4.39(3)

TABLE II. Autocorrelation times for the three-dimensional Gaussian model. Integrated autocorrelation times for $t_f/L^2 = 1/4 = 0.25$. The distance between two subsequent measurements is $t_{\mathrm{meas}} = L^3/32 = t_f L/8$. For each L, 5×10^6 measurements are performed. The τ are given in units of t_{meas}.

L	$\tau_{\mathrm{int},1}$	$\tau_{\mathrm{int},2}$	$\tau_{\mathrm{int},3}$	$\tau_{\mathrm{int},4}$
16	17.32(15)	13.00(11)	7.85(9)	7.59(5)
32	16.71(14)	10.53(9)	7.56(8)	7.12(5)
64	17.08(14)	9.02(8)	7.45(8)	7.04(5)
128	16.82(14)	7.75(5)	7.29(7)	6.87(5)

2. Three dimensions

Next we simulated the three-dimensional Gaussian model. We simulated the linear lattice sizes $L = 16$, 32, 64, and 128 for various values of t_f. Similar to the two-dimensional case, we use the masses $m = 0.6$, 0.3, 0.15, and 0.075. In Fig. 1 on the right, we give the integrated autocorrelation time of the observable O_3. Also in three dimensions it turns out that the curves for the different L fall approximately on top of each other for plotting τ_{int} as a function of t_f/L^2. Since we have taken $L \propto m^{-1}$, it is likely that actually $m^2 t_f$ should be kept constant.

As for two dimensions, we have measured the number of events per t_{MC}. We get here ≈1.38. For $t_f/L^2 = 1/4 = 0.25$ we performed simulations with 5×10^6 measurements. The results of these simulations are summarized in Table II. In contrast to two dimensions, the autocorrelation times are slightly decreasing with increasing lattice size (decreasing mass). Hence slowing down is eliminated.

B. The $O(3)$-invariant nonlinear σ-model in two dimensions

We performed updates in the (0, 1), (0, 2) and (1, 2) planes of the spins in a fixed order. For each of the planes, we used $t_f = L^2$. We simulated the model at $\beta = 1.4$, 1.5, 1.6, 1.7, and 1.8 on lattices of the linear size $L = 68$, 110, 190, 346, and 646, respectively. Following Ref. [13] the correlation length is $\xi = 6.90(1)$, 11.09(2), 19.07(6), 34.57(7), and 64.78(15) at these values of β, respectively. Consistent results are given in Ref. [19].

Our results for the observables

$$E = \frac{1}{L^2}\sum_{x,\mu} \vec{s}_x \cdot \vec{s}_{x+\hat{\mu}}, \qquad \chi = \frac{1}{L^2}\left(\sum\nolimits_x \vec{s}_x\right)^2 \quad (25)$$

and the corresponding integrated autocorrelation times are summarized in Table III. The integrated autocorrelation times are given in units of three XY embeddings. In the last column we give the average number of events divided by the number of sites for one cycle of three XY embeddings. The number slowly increases with increasing lattice size. This has to be taken into account when judging the performance of the algorithm. The fact that the number is close to 1 in all cases means that in one cycle, each site is touched roughly once.

In the case of the energy density E, the integrated autocorrelation time even decreases with increasing correlation length ξ. On the other hand, for the susceptibility χ, the integrated autocorrelation time is increasing with increasing ξ. Fitting with the ansatz $\tau_{\mathrm{int},\chi} N_{ev}/L^2 = c\xi^z$ gives $z = 0.386(5)$ and $\chi^2/\mathrm{d.o.f.} = 7.74$, taking all data into account and $z = 0.367(7)$ and $\chi^2/\mathrm{d.o.f.} = 4.93$, taking all data with $\beta \geq 1.5$ into account. On the other hand, fitting with the ansatz $\tau_{\mathrm{int},\chi} N_{ev}/L^2 = a + c\ln(\xi)$ gives $c = 0.74(1)$ and $\chi^2/\mathrm{d.o.f.} = 7.82$, taking all data into account and $c = 0.80(2)$ and $\chi^2/\mathrm{d.o.f.} = 1.33$, taking all data with $\beta \geq 1.5$ into account. The logarithmic increase of the integrated autocorrelation time is favored by the fits. In any case, slowing down is essentially eliminated, since also the fits with a power law ansatz result in small values for z.

Note that in Ref. [3] slowing down with $z \approx 1$ has been observed for the $O(3)$-invariant model on the three-dimensional simple cubic lattice. Furthermore one should notice that the single cluster algorithm [13] is more efficient than the event-chain Monte Carlo algorithm and on top of that provides variance reduced estimators for various observables.

C. The $SU(3) \times SU(3)$-invariant principal chiral model

For an update sequence of the length t_f a fixed random rotation matrix R is used. Then for the next update sequence a new R is generated.

TABLE III. Our results for the $O(3)$-invariant nonlinear σ-model on the square lattice.

β	E	$\tau_{\mathrm{int},E}$	χ	$\tau_{\mathrm{int},\chi}$	N_{ev}/L^2
1.4	1.124340(23)	1.476(10)	78.65(10)	1.358(9)	0.954050(12)
1.5	1.203250(14)	1.507(11)	176.67(25)	1.597(13)	0.995291(9)
1.6	1.271421(7)	1.503(12)	447.42(69)	1.937(17)	1.033550(6)
1.7	1.3284761(35)	1.376(12)	1273.9(2.1)	2.306(26)	1.070078(4)
1.8	1.3758734(16)	1.256(11)	3841.2(7.0)	2.739(31)	1.105618(3)

TABLE IV. Parameters of the runs for the principal chiral model and our results for the correlation length ξ_F.

β	L	ξ_F	L/ξ_F
1.5	32	3.01629(4)	10.6
1.65	64	5.4549(2)	11.8
1.825	128	11.679(6)	11.0
1.985	256	23.41(5)	10.9

We measure the following observables, constructed from fundamental and adjoint correlation functions [16]

$$G_F(x-y) = \langle \mathrm{tr}(U_x^\dagger U_y) \rangle;$$
$$G_A(x-y) = \langle |\mathrm{tr}(U_x^\dagger U_y)|^2 \rangle - 1. \qquad (26)$$

The energies E_F and E_A and the susceptibilities χ_F and χ_A

$$E_F = \frac{1}{N} G_F(1,0); \qquad E_A = \frac{1}{N^2-1} G_A(1,0) \qquad (27)$$

$$\chi_F = \sum_{x\in\Lambda} G_F(x); \qquad \chi_A = \sum_{x\in\Lambda} G_A(x) \qquad (28)$$

as well as the second moment correlation length

$$\xi_F = \frac{(\chi_F/F_F - 1)^{1/2}}{2\sin(\pi/L)}; \qquad F_F = \sum_x e^{2\pi i x_0/L} G_F(x). \qquad (29)$$

In Table IV we give the basic parameters of our runs and our result for the correlation length ξ_F. We have chosen the linear lattice size L such that $L/\xi_F \approx 11$ for all values of β. We performed a large number of simulations, varying the length t_f of the update sequence. We find that the results from various runs are compatible and agree also with results from the literature [16,17], confirming the correctness of our implementation of the algorithm.

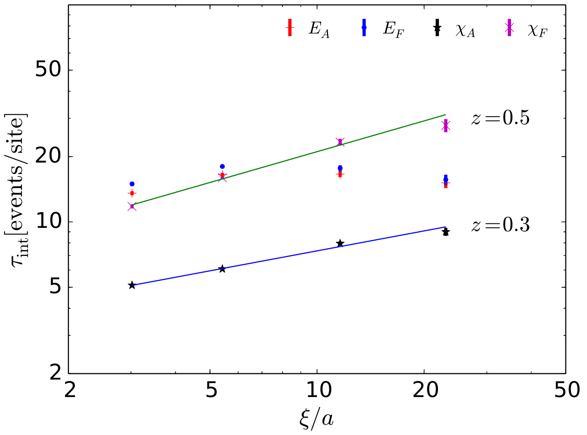

FIG. 3. Scaling of the integrated autocorrelation times with the correlation length. We observe dynamical critical exponents z well below 1 in all cases.

In this particular study, we also measure in time intervals which are smaller than t_f in order to reveal the scaling behavior with long update sequences. Our estimates of the integrated autocorrelation times are shown in Fig. 2. We give both the autocorrelation time and the length t_f of the update sequence normalized by the average number of events divided by the number of sites L^2. We observe a rather weak dependence on the length t_f of the update sequence, with a shallow minimum. In particular for sequences with a length t_f such that 0.1 up to 1 events per site occur, we see very little variation.

Next we study in more detail how the autocorrelation times scale with the correlation length ξ_F. To this end, for each value of β, we determine t_f at which the integrated autocorrelation time is minimal. In Fig. 3, the corresponding minimal integrated autocorrelation times for four

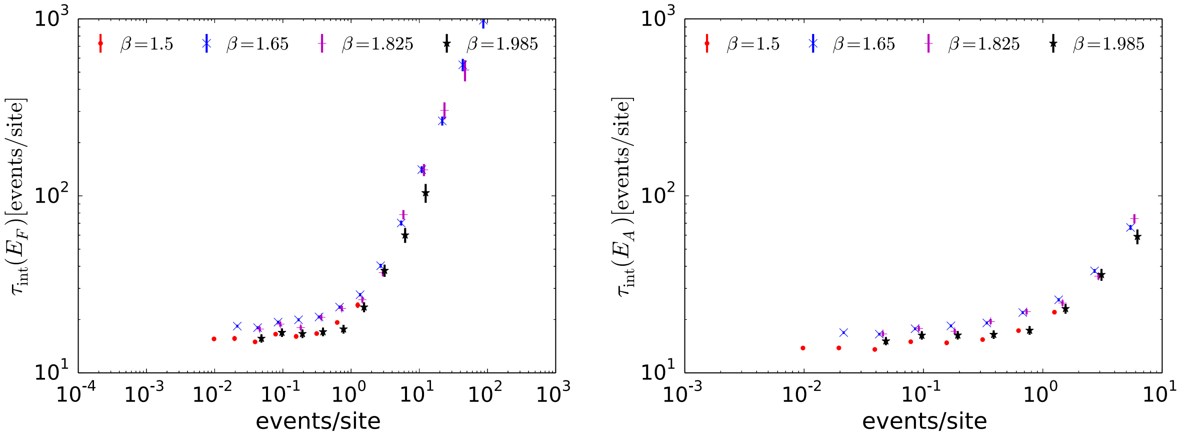

FIG. 2. Integrated autocorrelation time of the susceptibility as a function of the average number of events within one update sequence divided by the lattice volume. One the left/right we give the results for the fundamental/adjoint representation. In both cases, we observe a weak dependence on the length of the sequence and a shallow minimum.

different observables are plotted against the correlation length.

The behavior of $\tau_{\mathrm{int},\chi_A}$ and $\tau_{\mathrm{int},\chi_F}$ can be described reasonably well by a power law behavior with $z \approx 0.3$ and 0.5, respectively. This behavior is indicated by the two solid lines in Fig. 3. In the case of the energies E_A and E_F, the integrated autocorrelation is even decreasing, going from $\beta = 1.65$ to 1.985. We conclude that the dynamical critical exponent z is well below 1 and hence critical slowing down is essentially eliminated.

V. CONCLUSIONS AND OUTLOOK

We applied the event-chain Monte Carlo algorithm [1–4] to asymptotically free models in two dimensions. In the literature, these models are considered as toy models of lattice QCD that allow for testing of algorithmic and theoretical ideas in a simplified setting. In a preliminary study, we applied the event-chain Monte Carlo to free field theory on a square and a simple cubic lattice. We find that critical slowing down is eliminated. This is quite astonishing for an algorithm, where local changes of the field variables are performed. Next we studied the two-dimensional XY-model. We do not discuss our results in detail. They corroborate the findings of [18]: At low temperatures, in the spin wave phase, slowing down is eliminated, which is consistent with our free field theory result. On the other hand, in the high temperature phase, the physics is governed by vortices. Here we find slowing down with $z \approx 2$, even though the amplitudes of the integrated autocorrelation times are considerably reduced compared with the local Metropolis algorithm. One should note however that for this model the over-relaxation algorithm [14] gives $z \approx 1$ in both phases and the cluster algorithm eliminates slowing down completely [13,20]. In order to apply the event-chain Monte algorithm to the $O(3)$-invariant nonlinear σ-model and the $SU(3) \times SU(3)$ chiral model, we use an embedding of the XY-model. In the case of the $O(3)$-invariant model this simply means that one picks out two of the three components of the field. During one sequence of updates only these two components are updated. For the next sequence a new pair of components is taken. In the case of the $SU(3) \times SU(3)$ principal chiral model the embedding is a bit more complicated. We follow Ref. [9] whose authors used such an embedding in the context of the multigrid algorithm. For details see Sec. III C.

We find for both the $O(3)$-invariant nonlinear σ-model and the $SU(3) \times SU(3)$ chiral model that critical slowing down is essentially eliminated. While this is certainly encouraging with regard to the application to lattice QCD, open questions remain. What should a coherent embedding of $U(1)$ variables in the gauge theory look like? Will the event-chain Monte Carlo algorithm speed up the decorrelation of topological objects that are present in QCD?

ACKNOWLEDGMENTS

M. H. thanks the Deutsche Forschungsgemeinschaft for support under Grant No. HA 3150/4-1.

APPENDIX A: PROOF OF STABILITY

We have to show that the infinitesimal update by ϵ preserves the target distribution

$$\pi(X) = \frac{1}{2VZ} \exp(-S[\phi]), \tag{A1}$$

where Z is the partition function of the lattice model and the factors 2 and V give the number of possible values of σ and $\tilde{x}$. Let us write the update probabilities defined in Sec. II as $T(Y \leftarrow X)$. Then stability means

$$\pi(Y) = \sum_X T(Y \leftarrow X) \pi(X), \tag{A2}$$

where $\sum_X$ is a shorthand for the integral over the field ϕ and the sums over $\tilde{x}$ and σ. As a first step, we list the configurations X that can end up in a given $Y = (\phi, \sigma, \tilde{x})$, after an update step.

(i) X_0 given by ϕ the same as for Y, σ is replaced by $-\sigma$ and the position of the walker $\tilde{x}$ remains the same.

(ii) X_i given by ϕ is the same as for Y, σ is the same as for Y and $\tilde{x} = y_i$, meaning that the walker is hopping to its interaction partner y_i.

(iii) X_{n+1} given by $\phi_{\tilde{x}} - \sigma\epsilon$, ϕ_x with $x \neq \tilde{x}$ are the same as for Y, σ and $\tilde{x}$ keep their value.

Hence Eq. (A2) can be written as

$$\pi(Y) = \sum_{i=0}^{n+1} T(Y \leftarrow X_i) \pi(X_i). \tag{A3}$$

Note that the probability density of X_i for $i \leq n$ is the same as that of Y, since the field ϕ is the same. The probability density of X_{n+1} is

$$\begin{aligned} \pi(X_{n+1}) &= \pi(Y) \exp\left[-\sum_i [s_{i,x}(\phi_{\tilde{x}} - \sigma\epsilon) - s_{i,x}(\phi_{\tilde{x}})]\right] \\ &= \pi(Y)\left[1 + \sigma \sum_{i=0}^{n} \frac{\partial s_{i,\tilde{x}}}{\partial \phi_{\tilde{x}}} \epsilon + O(\epsilon^2)\right]. \end{aligned} \tag{A4}$$

Now let us work out the transition probabilities,

$$\begin{aligned} T(Y \leftarrow X_i) &= 1 - p_i \\ &= \max\left[0, -\sigma \frac{\partial s_{i,\tilde{x}}}{\partial \phi_{\tilde{x}}} \epsilon\right] + O(\epsilon^2) \quad \text{for } i \leq n. \end{aligned} \tag{A5}$$

Note that for $i = 0$ the sign of σ changes, while for $1 \leq i \leq n$ the hopping of the walker along with Eq. (3) produces a minus sign. Finally

$$T(Y\leftarrow X_{n+1})=\prod_{i=0}^{n}p_i=\prod_{i=0}^{n}\min\left[1,1-\sigma\frac{\partial s_{i,\tilde{x}}}{\partial\phi_{\tilde{x}}}\epsilon\right]+O(\epsilon^2)$$
$$=1+\sum_{i=0}^{n}\min\left[0,-\sigma\frac{\partial s_{i,\tilde{x}}}{\partial\phi_{\tilde{x}}}\epsilon\right]+O(\epsilon^2). \quad (A6)$$

Now we can put things together,

$$\sum_{i=0}^{n}T(Y\leftarrow X_i)\pi(X_i)+T(Y\leftarrow X_{n+1})\pi(X_{n+1})$$
$$=\pi(Y)\left(\sum_{i=0}^{n}\max\left[0,-\sigma\frac{\partial s_{i,\tilde{x}}}{\partial\phi_{\tilde{x}}}\epsilon\right]+O(\epsilon^2)\right)$$
$$+\pi(Y)\left(1+\sigma\sum_{i=0}^{n}\frac{\partial s_{i,\tilde{x}}}{\partial\phi_{\tilde{x}}}\epsilon+O(\epsilon^2)\right)$$
$$\times\left(1+\sum_{i=0}^{n}\min\left[0,-\sigma\frac{\partial s_{i,\tilde{x}}}{\partial\phi_{\tilde{x}}}\epsilon\right]+O(\epsilon^2)\right)$$
$$=\pi(Y)[1+O(\epsilon^2)]. \quad (A7)$$

Contributions $O(\epsilon)$ exactly cancel each other.

APPENDIX B: AUTOCORRELATION TIMES

The autocorrelation function of an observable A is given by

$$\rho_A(t)=\frac{\langle A_iA_{i+t}\rangle-\langle A\rangle^2}{\langle A^2\rangle-\langle A\rangle^2}, \quad (B1)$$

where the index of A gives the position in the Markov chain. The modulus of the autocorrelation function is bounded from above by an exponentially decaying function. Following Ref. [21] we define

$$\tau_{\mathrm{exp},A}=\lim_{t\to\infty}\sup\frac{t}{-\log(|\rho_A(t)|)} \quad (B2)$$

and the exponential autocorrelation time as

$$\tau_{\mathrm{exp}}=\mathrm{sup}_A\tau_{\mathrm{exp},A}, \quad (B3)$$

which characterizes the Markov chain. Since τ_{exp} is hard to compute, we determine integrated autocorrelation times. The integrated autocorrelation time of the observable A is given by

$$\tau_{\mathrm{int},A}=0.5+\sum_{t=1}^{\infty}\rho_A(t). \quad (B4)$$

The statistical error of the estimate of $\langle A\rangle$ is

$$\epsilon_A=\sqrt{\frac{2\tau_{\mathrm{int},A}}{n}\sigma_A^2}, \quad (B5)$$

where n is the number of measurements and

$$\sigma_A^2=\langle A^2\rangle-\langle A\rangle^2 \quad (B6)$$

is the variance. Computing τ_A, we use the estimate of $\rho_A(t)$ obtained from our simulation. Therefore the summation in Eq. (B4) has to be truncated at some finite t_{max}. Since $\rho_A(t)$ is falling off exponentially at large distances, the relative statistical error becomes large at large distances. Therefore it is mandatory to truncate the summation at some point that is typically much smaller than the total length of the simulation. In the literature one can find various recommendations on how this upper bound should be chosen. Sokal [21] suggests to take $t_{\mathrm{max}}\approx M\tau_{\mathrm{int}}$, where the value of M should be at least 6.

[1] E. P. Bernard, W. Krauth, and D. B. Wilson, Event-chain Monte Carlo algorithms for hard-sphere systems, Phys. Rev. E **80**, 056704 (2009).

[2] M. Michel, J. Mayer, and W. Krauth, Event-chain Monte Carlo for classical continuous spin models, Europhys. Lett. **112**, 20003 (2015).

[3] Y. Nishikawa, M. Michel, W. Krauth, and K. Hukushima, Event-chain algorithm for the Heisenberg model: Evidence for $z\simeq 1$ dynamic scaling, Phys. Rev. E **92**, 063306 (2015).

[4] S. Kapfer and W. Krauth, Irreversible Local Markov Chains with Rapid Convergence towards Equilibium, Phys. Rev. Lett. **119**, 240603 (2017).

[5] R. H. Swendsen and J.-S. Wang, Nonuniversal Critical Dynamics in Monte Carlo Simulations, Phys. Rev. Lett. **58**, 86 (1987).

[6] U. Wolff, Collective Monte Carlo Updating for Spin Systems, Phys. Rev. Lett. **62**, 361 (1989).

[7] J. Goodman and A. D. Sokal, Multigrid Monte Carlo method. Conceptual foundations, Phys. Rev. D **40**, 2035 (1989).

[8] R. C. Brower and P. Tamayo, Embedded Dynamics for φ^4 Theory, Phys. Rev. Lett. **62**, 1087 (1989).

[9] T. Mendes, A. Pelissetto, and A. D. Sokal, Multigrid Monte Carlo via XY embedding. General theory and

two-dimensional $O(N)$-symmetric non-linear σ-models, Nucl. Phys. **B477**, 203 (1996).

[10] M. Hasenbusch and K. Pinn, Computing the roughening transition of Ising and solid-on-solid models by BCSOS model matching, J. Phys. A **30**, 63 (1997).

[11] M. Hasenbusch, The Binder cumulant at the Kosterlitz-Thouless transition, J. Stat. Mech. Theor. Exp. **2008**, P08003 (2008).

[12] J. Balog, M. Niedermaier, F. Niedermayer, A. Patrascioiu, E. Seiler, and P. Weisz, Comparison of the O(3) bootstrap σ model with the lattice regularization at low energies, Phys. Rev. D **60**, 094508 (1999).

[13] U. Wolff, Asymptotic freedom and mass generation in the O(3) nonlinear σ-model, Nucl. Phys. **B334**, 581 (1990).

[14] J. Apostolakis, C. F. Baillie, and G. C. Fox, Investigation of the two-dimensional O(3) model using the overrelaxation algorithm, Phys. Rev. D **43**, 2687 (1991).

[15] U. Wolff, Simulating the all-order strong coupling expansion III: $O(N)$ sigma/loop models, Nucl. Phys. **B824**, 254 (2010); Erratum **B834**, 395(E) (2010).

[16] G. Mana, A. Pelissetto, and A. D. Sokal, Multigrid Monte Carlo via XY embedding. II. Two-dimensional SU(3) principal chiral model, Phys. Rev. D **55**, 3674 (1997).

[17] M. Hasenbusch and S. Meyer, Multigrid acceleration for asymptotically free theories, Phys. Rev. Lett. **68**, 435 (1992).

[18] Z. Lei and W. Krauth, Irreversible Markov chains in spin models: Topological excitations, Europhys. Lett. **121**, 10008 (2018).

[19] J. Balog, M. Niedermaier, F. Niedermayer, A. Patrascioiu, E. Seiler, and P. Weisz, Comparison of the O(3) bootstrap σ model with the lattice regularization at low energies, Phys. Rev. D **60**, 094508 (1999).

[20] U. Wolff, Collective Monte Carlo updating in a high precision study of the x-y model, Nucl. Phys. **B322**, 759 (1989).

[21] A. D. Sokal, Monte Carlo Methods In Statistical Mechanics: Foundations and new Algorithms, *Cours de Troisième Cycle* (Lausanne, Switzerland, 1989); N. Madras and A. D. Sokal, The pivot algorithm: A highly efficient Monte Carlo method for the self-avoiding walk, J. Stat. Phys. **50**, 109 (1988).