Efficient predecision scheme for Metropolis Monte Carlo simulation of long-range interacting lattice systems

Fabio Müller[1, *] and Wolfhard Janke[1, †]

[1]*Institut für Theoretische Physik, Universität Leipzig, IPF 231101, 04081 Leipzig, Germany*

(Dated: August 14, 2025)

We propose a fast and general predecision scheme for Metropolis Monte Carlo simulation of d-dimensional long-range interacting lattice models. For potentials of the form $V(r) = r^{-d-\sigma}$, this reduces the computational complexity from $O\left(N^2\right)$ to $O\left(N^{2-\sigma/d}\right)$ for $\sigma < d$ and to $O\left(N\right)$ for $\sigma > d$, respectively. The algorithm is implemented and tested for several O(n) spin models ranging from the Ising over the XY to the Edwards-Anderson spin-glass model. With the same random number sequence it produces exactly the same Markov chain as a simulation with explicit summation of all terms in the Hamiltonian. Due to its generality, its simplicity, and its reduced computational complexity it has the potential to find broad application and thus lead to a deeper understanding of the role of long-range interactions in the physics of lattice models, especially in nonequilibrium settings.

The Metropolis Monte Carlo algorithm [1] can be seen as the workhorse for the investigation of lattice systems, allowing the study of both equilibrium properties as well as nonequilibrium processes. One broad class of models are vector spin (O(n)) models which in various definitions capture diverse physical behavior ranging from first- and second-order phase transitions to infinite-order BKT-type transitions or even glassy behavior. Also from a nonequilibrium perspective there is a broad range of different processes which can be studied with the Metropolis algorithm and which recently have attracted significant attention, such as field-driven hysteresis, phase ordering, Kibble-Zurek mechanism, Mpemba effect, etc. From a computational point of view, these physical aspects are usually studied under the restriction to short-range interacting systems since the conventional simulation of a long-range interacting system with N constituents comes along with an $O(N^2)$ complexity. This fact has severely impeded simulation studies and only in very recent times algorithmic advances have alleviated this problem, leading to a new wave of important insights into the nonequilibrium physics of long-range interacting systems [2–6]. One keystone for this development has been our recent fast and hierarchical Metropolis algorithm [7] which, however, comes with a fairly elevated technical overhead constituting a barrier to its broad adoption.

In this Letter we propose a new, more streamlined algorithm. Opposed to the fast, hierarchical algorithm [7], it is even applicable for systems with random signs in the interaction. Regarding the technical requirements for its implementation, the new algorithm is on par with the conventional Metropolis algorithm, while still providing a significant reduction of the computational complexity. Since the algorithm, using the same random-number sequence, reproduces exactly the same Markov chain as the Metropolis algorithm with the exact evaluation of the energy, all existing physical knowledge about the systems under study can be carried over one-to-one. While this property makes it particularly useful for the investigation of nonequilibrium processes, the algorithm can, of course, also be applied to equilibrium systems. The property of reproducing the Metropolis criterion, the simplicity of its implementation, and the coincident enormous reduction of the computational burden render it a very promising candidate for broad application in studies of a vast spectrum of phenomena enrooted in long-range interacting systems.

As prototypical models we consider three different variants of long-range O(n) spin models which all have in common the Hamiltonian

$$\mathcal{H} = -\frac{1}{2} \sum_i \sum_{j \neq i} J_{i,j} \mathbf{s}_i \mathbf{s}_j, \tag{1}$$

where the spins $\mathbf{s}_i$ are n-dimensional unit vectors, including the Ising model for $n = 1$. The interactions are given by

$$|J_{i,j}| = r(i,j)^{-d-\sigma}, \tag{2}$$

where $r(i,j)$ denotes the Euclidean distance between $\mathbf{s}_i$ and $\mathbf{s}_j$, d is the spatial dimension, and $\sigma > 0$ is a tuneable parameter determining the decay of the interactions. The absolute value $|J_{i,j}|$ is taken to account for the possibility of antiferromagnetic interactions which, for instance, appear in the long-range Edwards-Anderson spin-glass model which we consider as one example. As in our Ref. [7] we implement periodic boundary conditions using Ewald summation [8], resulting in a modified set of couplings which is determined once at the beginning of the simulation and then used together with the simple minimum-image convention.

To cover a broad range of features of the model class defined by Eq. (1), we consider the following cases: i) the paradigmatic ferromagnetic long-range Ising model (LRIM), ii) the long-range ferromagnetic XY model (LRXYM) being the most simple vector spin model,

and iii) the long-range Edwards-Anderson spin glass (LREAM), as prototypical model where the sign of the interaction is a quenched random variable determined individually for each pair of spins $\mathbf{s}_i$ and $\mathbf{s}_j$ at the beginning of the simulation. The former two systems fulfill translation invariance such that a single set of couplings $J_{0,k}$ can be used for the simulation. The latter one additionally needs a matrix $\iota_{i,j}$ containing the sign of the interaction for each pair of spins.

In its traditional implementation, the Metropolis algorithm is based on a proposal-acceptance scheme. For the system which is currently in a microstate with energy E^{old} a new microstate with energy E^{new} is proposed. The new microstate is accepted if

$$\rho \leq \exp(-\beta \Delta E), \tag{3}$$

where $\rho \in [0,1)$ is a random number, β the inverse temperature, and $\Delta E = E^{\text{new}} - E^{\text{old}}$ is the energy change associated with the proposed update. Usually, for spin systems an update only concerns a single spin and hence the calculation of ΔE in a model with only nearest-neighbor interactions comprises a constant, very limited number of terms. For a long-range O(n) model the calculation of ΔE resulting from the update of a single spin $\mathbf{s}_i$ reads

$$\Delta E = E_i^{\text{new}} - E_i^{\text{old}} = -\Delta \mathbf{s}_i \sum_{j \neq i} J_{i,j} \mathbf{s}_j, \tag{4}$$

where $\Delta \mathbf{s}_i = \mathbf{s}_i^{\text{new}} - \mathbf{s}_i^{\text{old}}$ and the summation runs over all other spins in the system. Thus, it requires the summation over the whole system, resulting in the infamous $O(N^2)$ complexity per Monte Carlo sweep (MCS) consisting of a number of updates equal to the number of spins contained in the system. However, as shown in Refs. [7, 9] this problem can in many cases be circumvented. The full evaluation of (4) is avoided by solving (3) for ΔE

$$\Delta E \leq -\frac{\ln \rho}{\beta} \equiv E^{\text{th}}, \tag{5}$$

defining the threshold energy E^{th} up to which an update gets accepted. This allows decisions based upon strict bounds $\Delta E^{\text{min}} \leq \Delta E$ and $\Delta E^{\text{max}} \geq \Delta E$. If ΔE^{min} is larger than E^{th} the update is rejected, while it is accepted if ΔE^{max} is smaller than E^{th}. Depending on the circumstances, the decisions can be taken with an extremely reduced computational effort. In Ref. [7] we built upon this idea, establishing a generic algorithm for both lattice and off-lattice systems. This was achieved by leveraging the tree-structure of a hierarchical spatial decomposition of the simulation domain. Here, we propose a very simplified framework based on partial sums combined with a predecision scheme. To achieve this, we first introduce a permutation $\pi(j)$ in (4) which allows to rearrange the summation order to best fit our algorithmic purposes

(which in most cases is given by sorting $J_{i,j}$ in decreasing order with respect to their magnitude). Secondly we split the sum into two parts,

$$\Delta E = \underbrace{-\Delta \mathbf{s}_i \sum_{j=1}^{n} J_{i,\pi(j)} \mathbf{s}_{\pi(j)}}_{K \equiv \text{known part}} \underbrace{-\Delta \mathbf{s}_i \sum_{j=n+1}^{N-1} J_{i,\pi(j)} \mathbf{s}_{\pi(j)}}_{R \equiv \text{rest}}$$
$$= K(n) + R(n) \tag{6}$$

of which only the first sum $K(n)$ will be treated explicitly and thus is known exactly. For the treatment of $R(n)$ we introduce an additional quantity

$$U(n) = |\Delta \mathbf{s}_i| \sum_{j=n+1}^{N-1} \left| J_{i,\pi(j)} \right| = |\Delta \mathbf{s}_i| \left[J_i^{\text{int}} - \sum_{j=1}^{n} \left| J_{i,\pi(j)} \right| \right], \tag{7}$$

where

$$J_i^{\text{int}} = \sum_{j=1}^{N-1} |J_{i,j}|, \tag{8}$$

is the sum over the couplings between $\mathbf{s}_i$ and all other spins in the system. Since J_i^{int} are calculated once at the beginning of the simulation (in fact for systems with periodic boundary conditions and translation invariance there is only a single J_0^{int} which in the following will be simply denoted as J^{int}), $U(n)$ can be calculated alongside with $K(n)$ during the summation process. We note that for each summand in $R(n)$, $-|\Delta \mathbf{s}_i J_{i,k}| \leq -\Delta \mathbf{s}_i J_{i,k} \mathbf{s}_k \leq |\Delta \mathbf{s}_i J_{i,k}|$ (because the spins are unit vectors) and thus

$$-U(n) \leq R(n) \leq U(n). \tag{9}$$

Now, we can easily establish strict lower and upper bounds on (6)

$$\Delta E^{\text{min}}(n) = K(n) - U(n), \quad \text{and}$$
$$\Delta E^{\text{max}}(n) = K(n) + U(n), \tag{10}$$

which allow a predecision by prematurely terminating the sum at n_0 summation steps if either $\Delta E^{\text{min}}(n_0) > E_{\text{th}}$ (rejection) or $\Delta E^{\text{max}}(n_0) < E_{\text{th}}$ (acceptance).

Finally, it remains to choose an efficient permutation $\pi(j)$ such that $U(n)$ in Eq. (7) diminishes as efficiently as possible. The most intuitive approach, which we adopt here, is to pick a permutation that arranges $|J_{i,j}|$ in monotonously decreasing order to minimize $U(n)$ for given n [10].

The explicit algorithmic formulation for a system with translation symmetry contains the following preparatory steps

1. Calculate all interactions $J_{0,k}$ from the reference spin at $\mathbf{r} = \mathbf{0}$ to all other spins in the lattice.

2. Store these interactions alongside their corresponding relative position vectors $\mathbf{r}_k$.
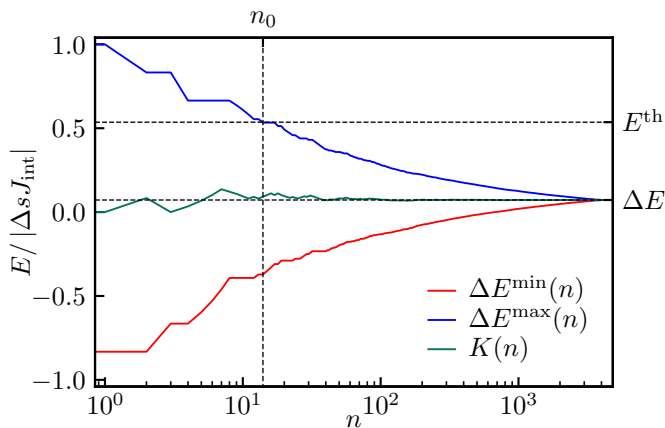
FIG. 1. Exemplary illustration of the predecision process for the long-range Ising model in $d = 2$ with $\sigma = 0.6$, $L = 64$, and $T = 20$. Due to the elevated temperature which is well above the critical temperature T_c the system is in a disordered state. The green line represents the exactly known part $K(n)$ which accounts for the actual sign of the individual interactions. The red and the blue line show the evolution of the upper and lower bound of $\Delta E^{\mathrm{min/max}}(n) = K(n) \mp U(n)$, respectively which by construction converge monotonously to the exact value of ΔE associated with the update. Already at $n_0 = 14$ (of $N - 1 = 64^2 - 1 = 4095$) summation steps the upper bound ΔE^{max} lies below the energy threshold E^{th} such that the update can be accepted with the explicit knowledge of only a tiny fraction of all interactions.

3. Sort this list in descending order for the interaction strength $J_{0,k}$.

4. Compute the integrated interaction $J^{\mathrm{int}} = \sum_{j=1}^{N-1} J_{0,j}$.

Since these steps are only performed once at the beginning of the simulation, they do not significantly contribute to the algorithm's execution time. However, the spins $\mathbf{s}_j$ now must be properly identified using $\mathbf{r}_j = \mathbf{r}_i + \mathbf{r}_k$ taking into account the minimum-image convention. Each update then consists of the following steps:

1. Randomly select a spin $\mathbf{s}_i$ to update.

2. Propose a new state $\mathbf{s}_i^{\mathrm{new}}$ (for Ising spins $s_i^{\mathrm{new}} = -s_i^{\mathrm{old}}$)

3. Randomly draw the threshold energy E^{th} (5).

4. If $E^{\mathrm{th}} > |\Delta \mathbf{s}_i| \, J_i^{\mathrm{int}}$ accept the update directly.

5. Continue summation until

   • Either $\Delta E^{\mathrm{max}}(n) < E^{\mathrm{th}}$: Accept update.

   • Or $\Delta E^{\mathrm{min}}(n) > E^{\mathrm{th}}$: Reject update.

In Fig. 1 we illustrate this procedure by visualizing an exemplary decision process. There, we plot $K(n)$ together with both $\Delta E^{\mathrm{min}}(n)$ and $\Delta E^{\mathrm{max}}(n)$. One clearly

sees how the energy bounds approach the true value of ΔE for growing index n. In this specific situation, the update can be accepted already after $n_0 = 14$ ($\approx 0.34\%$ of total number of spins) summation steps because the difference between the threshold energy E^{th} and the actual value ΔE associated with the spin flip is fairly large.

Next, we turn to a performance analysis. The common approach to assess the computational complexity via the scaling of wall-clock times of the algorithm has the major drawback that it can strongly depend on several environmental factors. These are primarily the hardware architecture and configuration of the compute node, and the workload that is running in parallel, which can influence boost clocks and cache availability. While one can try to exactly reproduce the environment in which the code runs, it will still be difficult to safely assess runtimes with an uncertainty of less than 5%. Hence, we here choose a different approach by measuring the average number of needed summation steps $\langle n_0 \rangle$ per update, yielding an architecture and environment independent benchmark.

The resulting scaling analyses are presented in Fig. 2. The two-dimensional LRIM and LRXYM are benchmarked at the respective transition temperatures, while for the two-dimensional LREAM we perform the simulations at $0.5 T_c(\sigma)$ of the LRIM. We have checked, that the choice of the simulation temperature does not influence the asymptotic algorithmic scaling, rendering the choice of the simulation temperature secondary for our analysis. For all three models (LRIM, LRXYM, and LREAM) a power-law dependence of the form

$$\langle n_0 \rangle(N) = C + A N^b, \tag{11}$$

can be fitted to the data, for $\sigma \neq 2$, yielding exponents

$$b = 1 - \sigma/2. \tag{12}$$

By going into more details of the algorithm's performance and considering the probability density $f(n_0)$ and its first moment $\langle n_0 \rangle = \sum_{n_0=0}^{N-1} n_0 f(n_0)$, we have theoretical arguments [11] that the 2 in (12) should more generally be replaced by the dimension d. This is demonstrated in the Appendix for the example of $d = 1$. Hence, for $\sigma = d$ we expect a logarithmic divergence being supported by the good description of the data with the logarithmic ansatz $\langle n_0 \rangle(N) = a \ln N + C$ employed for $\sigma = 2$ in Fig. 2(a-c) for all three considered models. For $\sigma > 2$, $\langle n_0 \rangle$ approaches a constant for large N, i.e., in this short-range-like regime the exponent b plays the role of a subleading correction. The fit intervals are chosen individually for each fit to yield a χ^2 per degree of freedom close to unity. The resulting fits for all models and values of σ in the chosen fit interval are shown as solid black lines while the dotted lines represent their continuation to smaller N which were excluded from the fit. We thus find that the computational complexity is largely model independent, while the prefactors are model-dependent, varying by factors < 2. For $\sigma \geq 1.5$, compared to a simulation with
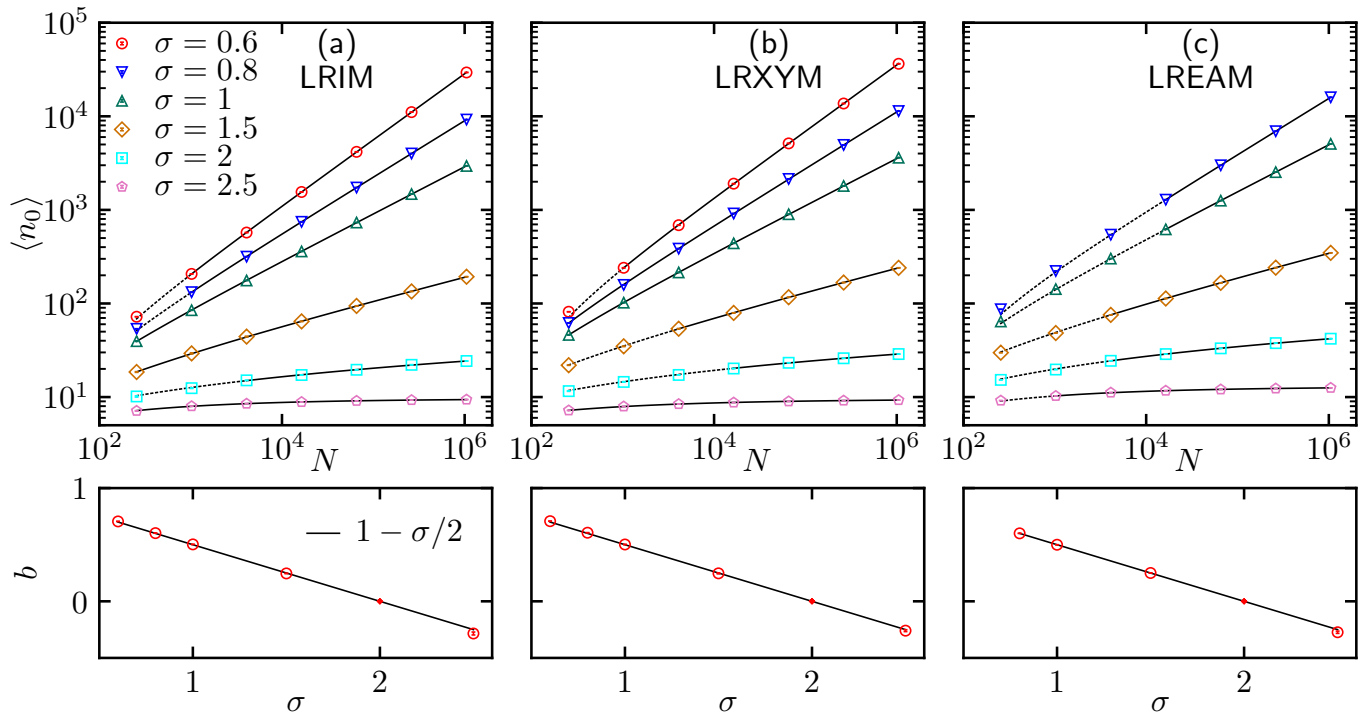
FIG. 2. Scaling analysis of the average number of needed summation steps $\langle n_0 \rangle$ for the (a) LRIM, (b) LRXYM, and (c) LREAM in $d = 2$ dimensions. The solid lines represent the fits to $\langle n_0 \rangle(N) = C + AN^b$ in the fitting interval to extract the asymptotic algorithmic scaling and the dotted lines their continuations to smaller N. The lower panel shows the resulting exponent b. For $\sigma = 2$ where we expect $b = 0$ a logarithmic ansatz is employed. For $\sigma > 2$ the computational effort per update becomes constant for $N \to \infty$.

a full calculation of the energy differences, already for $L = 1024$ we observe a reduction of the needed summation steps by factors between 3000 and 50 000 which due to the better scaling would further increase with growing system size. Hence, the algorithm enables simulations within days which would take decades with a naive implementation.

To conclude, we have presented a new algorithm for the exact treatment of long-range interaction for Metropolis Monte Carlo simulation of lattice systems. The algorithm can be thought of as a simulation with an adaptive cutoff which for each individual update is chosen just large enough to take the correct decision according to the Metropolis criterion. In a model-independent fashion, we observe a reduction of the computational complexity from $O(N^2)$ to $O(N^{2-\sigma/d})$ for $0 < \sigma < d$ and to $O(N)$ for $\sigma > d$, implying that for $\sigma > d$ even asymptotically it outperforms the technically significantly more challenging algorithm presented in Ref. [7]. For smaller values of σ close to d it can still be competitive up to very large system sizes, due to the algorithmic scaling remaining weak and the algorithmic prefactor being significantly smaller due to the reduced complexity of the algorithm.

We have not investigated any dynamical properties since the decisions of the Metropolis algorithm are exactly reproduced and thus also its dynamical properties.

The new method is very easy to implement and can easily be adapted to other use-cases. Finally, we also provide a working implementation in the Supplementary Material [12], such that the algorithm can easily be adopted and extended by other researchers in the field. Possible applications contain equilibrium studies of frustrated systems for which no efficient cluster algorithm is available such as, e.g., spin glasses, random-field models, or the dipolar Ising model [13]. Additionally, its features render the algorithm well suited for the investigation of nonequilibrium processes such as, e.g., phase-ordering kinetics [14], the Kibble-Zurek mechanism [15] or the Mpemba effect [16], and many others [17].

* fabio.mueller@itp.uni-leipzig.de
† wolfhard.janke@itp.uni-leipzig.de

[1] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculations by fast computing machines, J. Chem. Phys. 21, 1087 (1953).

[2] H. Christiansen, S. Majumder, and W. Janke, Phase or-

dering kinetics of the long-range Ising model, Phys. Rev. E **99**, 011301(R) (2019).

[3] H. Christiansen, S. Majumder, M. Henkel, and W. Janke, Aging in the long-range Ising model, Phys. Rev. Lett. **125**, 180601 (2020).

[4] H. Christiansen, S. Majumder, and W. Janke, Zero-temperature coarsening in the two-dimensional long-range Ising model, Phys. Rev. E **103**, 052122 (2021).

[5] F. Müller, H. Christiansen, and W. Janke, Phase-separation kinetics in the two-dimensional long-range Ising model, Phys. Rev. Lett. **129**, 240601 (2022).

[6] F. Müller, H. Christiansen, and W. Janke, Nonuniversality of aging during phase separation of the two-dimensional long-range Ising model, Phys. Rev. Lett. **133**, 237102 (2024).

[7] F. Müller, H. Christiansen, S. Schnabel, and W. Janke, Fast, hierarchical, and adaptive algorithm for Metropolis Monte Carlo simulations of long-range interacting systems, Phys. Rev. X **13**, 031006 (2023).

[8] P. Ewald, Die Berechnung optischer und elektrostatischer Gitterpotentiale, Ann. Phys. (Berl.) **369**, 253 (1921).

[9] S. Schnabel and W. Janke, Accelerating polymer simulation by means of tree data-structures and a parsimonious Metropolis algorithm, Comput. Phys. Commun. **256**, 107414 (2020).

[10] An imperfect ordering in $|J_{i,j}|$, where subsequences are additionally ordered to minimize cache misses, can lead to further improved wall-clock times.

[11] F. Müller and W. Janke, Analysis of runtime distributions in predecision Metropolis Monte Carlo, in preparation.

[12] See Supplemental Material at {will be added} where a full implementation of the algorithm in the Julia programming language is provided.

[13] K. De'Bell, A. B. MacIsaac, and J. P. Whitehead, Dipolar effects in magnetic thin films and quasi-two-dimensional systems, Rev. Mod. Phys. **72**, 225 (2000).

[14] A. J. Bray and A. D. Rutenberg, Growth laws for phase ordering, Phys. Rev. E **49**, R27 (1994).

[15] K. Du, X. Fang, C. Won, C. De, F.-T. Huang, W. Xu, H. You, F. J. Gómez-Ruiz, A. del Campo, and S.-W. Cheong, Kibble-Zurek mechanism of Ising domains, Nature Physics **19**, 1495 (2023).

[16] G. Teza, J. Bechhoefer, A. Lasanta, O. Raz, and M. Vucelja, Speedups in nonequilibrium thermal relaxation: Mpemba and related effects (2025), arXiv:2502.01758 [cond-mat.stat-mech].

[17] G. Ódor, Universality classes in nonequilibrium lattice systems, Rev. Mod. Phys. **76**, 663 (2004).

End Matter: Algorithmic scaling in $d = 1$

In the main text we have stated that the denominator 2 in (12) should actually read d. To corroborate on this statement we additionally perform a scaling analysis of $\langle n_0 \rangle$ for the LRIM in $d = 1$. As for the $d = 2$ LRIM the simulations are again performed at $T_c(\sigma)$. Since for $\sigma > 1$ the transition temperature is zero, the system in equilibrium is always in the ground state and the simulations become trivial. Therefore, we here restrict σ to $\sigma \leq 1$. As demonstrated in Fig. 3, the exponents in the scaling
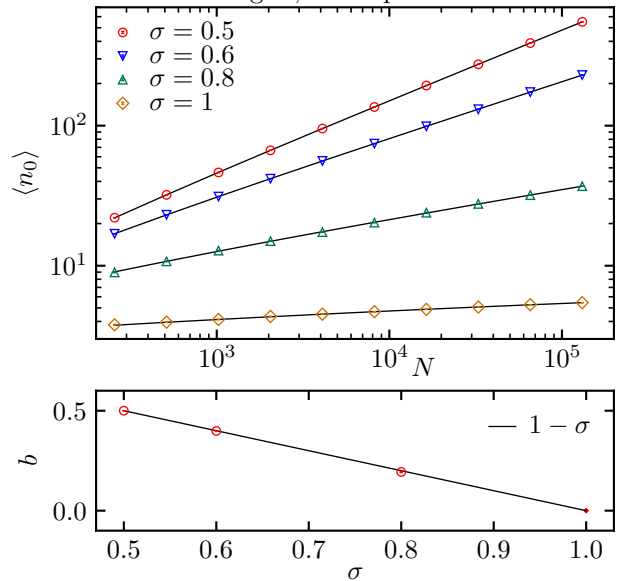

FIG. 3. Scaling analysis of the average number of needed summation steps $\langle n_0 \rangle$ for the $d = 1$ LRIM. The solid lines represent the fits to $\langle n_0 \rangle(N) = C + AN^b$ which in this case are performed over the full range of system sizes. The lower panel shows the exponent b of the power-law ansatz describing the asymptotic growth of $\langle n_0 \rangle$. For $\sigma = 1 = d$ where we expect $b = 0$ a logarithmic ansatz is taken.

ansatz (11) follow the relation $b = 1 - \sigma$ confirming that the general expression for b should read $b = 1 - \sigma/d$.