PHYSICAL REVIEW E **93**, 061301(R) (2016)

Minimal energy ensemble Monte Carlo algorithm for the partition function of fermions coupled to classical fields

Przemysław R. Grzybowski,[1] Łukasz Czekaj,[2] Mariusz Nogala,[1] Adam Ścibior,[1] and Ravindra W. Chhajlany[1,3]
[1]*Faculty of Physics, Adam Mickiewicz University in Poznań, Umultowska 85, 61-614 Poznań, Poland*
[2]*National Quantum Information Centre of Gdańsk, 81-824 Sopot, Poland*
[3]*ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain*
(Received 7 August 2015; revised manuscript received 25 January 2016; published 20 June 2016)

Models of noninteracting fermions coupled to auxiliary classical fields are relevant to the understanding of a wide variety of problems in many-body physics, e.g., the description of manganites, diluted magnetic semiconductors, or strongly interacting electrons on lattices. We present a flat-histogram Monte Carlo algorithm that simulates a statistical ensemble that allows one to directly acquire the partition function at all temperatures for such systems. The defining feature of the algorithm is that it utilizes the complete thermodynamic information from the full energy spectrum of noninteracting fermions available during sampling of the configuration space of the classical fields. We benchmark the method for the classical Ising and Potts models in two dimensions, as well as the Falicov-Kimball model describing itinerant electrons interacting with heavy ions.

DOI: 10.1103/PhysRevE.93.061301

Introduction. Bilinear Hamiltonians of lattice fermions coupled to classical degrees of freedom (continuous or discrete) are ubiquitous in contemporary many-body physics. These often arise as suitable approximations in the description of systems where many different degrees of freedom contrive to yield complex and interesting physics. In these cases, some subsystem is treated classically. e.g., localized spins in double-exchange models [1–3], models of Mn-doped (III and IV) semiconductors [4], the Ising t-J model [5], adiabatic phonons in polaron models [6–9], or one species of fermions in the Falicov-Kimball model [9–11]. Exactly solvable models can also take this form, such as the seminal honeycomb lattice Kitaev model [12], where interacting spins are mapped onto Majorana fermions coupled to static gauge fields. More generally, auxiliary field methods, e.g., based on the Hubbard-Stratonovich transformation, allow one to decouple fermion-fermion or fermion-boson interactions; the fields are then treated classically in conjunction with the application of the Suzuki-Trotter decomposition [13,14] (see [15] for a recent approximate scheme with static fields).

The simplicity of the form of such models belies their complexity. Although obtaining eigenstates for fixed configurations of classical fields is computationally easy, summing over the exponential number of such configurations to obtain thermodynamic properties is notoriously expensive. An obvious approach towards this problem is via Monte Carlo (MC) sampling [16]. Summing over the fermion states for a fixed classical field configuration yields the conditional (grand) partition function. This quantity, at fixed temperature, serves as the Metropolis weight [17] for performing a random walk in the space of classical field configurations [18]. The serious bottleneck in these simulations is the repeated performance of the fermionic trace, and so exact diagonalization (ED) of the free fermion system, for executing the walk. Hence various improvements have been proposed to optimize the reevaluation of the weight: moment expansion of the fermion density of states by Chebyshev polynomials [19,20], low-rank matrix updates [21], or Green's functions [22] Chebyshev expansion. On the other hand, it seems naturally essential to optimize the extraction of information at each MC step. Indeed, while the expensive ED yields the conditional partition function at all temperatures, these are completely discarded by using only single-temperature data in the above approaches. Here we introduce an algorithm that fully exploits the thermodynamic information available at each diagonalization step in MC simulations of free fermions coupled to classical fields (FCCFs).

The paradigmatic Metropolis algorithm [17] suffers from critical slowing down at continuous phase transitions and prolonged trapping in metastable states at discontinuous transitions. These can be overcome to a large extent using cluster update schemes [23–28] or sampling extended ensembles [29–32] such as Wang-Landau sampling of the density of states [33,34]. While these approaches have been used in the study of classical and quantum systems, FCCFs hold their own system-specific challenges rendering such applications difficult. In particular, the effective Hamiltonian (corresponding to the energies in the Metropolis weight) of the classical fields in general contains temperature-dependent long-range multiparticle interactions. Often, molecular dynamics or hybrid Monte Carlo methods using Langevin's equation [35–38] are better suited than standard MC simulations of such Hamiltonians, although the acceptance rate in these simulations crucially depends on the quality of the approximate action. Instead, we sample an extended ensemble, bringing the advantages of Wang-Landau-like sampling to FCCFs.

Problem and method. We first generalize our considerations: a system is biseparable if it may be separated into two subsystems A and B such that for a given state of A all states of the system B can be efficiently summed over to obtain the conditional (grand) partition function $Z(\beta|A)$ or equivalently the conditional free energy (grand potential) $F(\beta|A)$. This definition covers both FCCFs (where subsystem A is classical) and many classical models (e.g., the Ising model on a bipartite lattice where A and B are the spins on the two sublattices, respectively). For biseparable systems, the partition function

is decomposed as

$$Z(\beta) = \mathrm{Tr}\exp(-\beta H) = \sum_A \left(\sum_B \exp(-\beta H_{B|A}) \right)$$
$$= \sum_A Z(\beta|A) = \sum_A \exp[-\beta F(\beta|A)], \quad (1)$$

i.e., the partition function $Z(\beta)$ is obtained by averaging the conditional partition functions over all configurations attainable in the system A. Notice that once the conditional energy spectrum is obtained, complete thermodynamic information associated with the exponentially large number of configurations of B encoded in $\exp[-\beta F(\beta|A)]$ for all inverse temperatures β becomes potentially available at each simulation step.

Our problem can be stated as follows. Standard MC sampling for FCCF systems consists of walking, at fixed temperature, between different configurations of A with with Metropolis weights $\exp[-\beta F(\beta|A)]$. Here we aim to obtain the entire partition function (1) by acquiring, at a given simulation step, the conditional partition function $Z(\beta|A)$ for *all* arguments β from the information (full spectrum) abundantly available for each fixed configuration on A. In principle, this may also be achieved by the parallel tempering algorithm [32] (to efficiently sample configurational space of A) and a reweighting procedure, which, however, requires a well chosen set of temperature intervals. Instead, in our method, we perform a random walk directly in the configurational space of A without referring to any specific temperature. The basic challenge is to obtain an appropriate importance sampling scheme over the exponential number of configurations of A.

The key to any thermodynamic MC simulation lies in efficiently sweeping energy space. A simple observation is that, for most systems, only a few configurations of A lead to the spectrum of system energies containing the ground-state energy. Such configurations must obviously be effectively found by the importance sampling scheme. Moreover, the conditional densities of states (DOSs) $\rho_A(E)$ of energy spectra of typical configurations of A are expected to differ most also in their lower range. So a key discriminant for configurations of A is the minimal energy attainable by the system at a given configuration, which we denote by $\varepsilon_{\min}(A)$. We consider two configurations of A to be in the same class if they have equal $\varepsilon_{\min}$ values.

The principle that we identify and implement to acquire the partition function is that all minimal energy classes be visited by the algorithm. For this, notice that the minimal attainable energies $\varepsilon_{\min}$ can be associated with a density of states $\tilde{\rho}(\varepsilon_{\min})$, which enumerates the number of configurations of subsystem A attaining a given minimal energy. We emphasize that this auxiliary DOS is distinct from the true DOS of the system and does not determine the latter.

Our algorithm then consists of two separate sampling stages associated with first generating the weight distribution for importance sampling and then using it for acquiring thermodynamic information about the system. (i) The auxiliary DOS $\tilde{\rho}(\varepsilon_{\min})$ is readily obtained via a Wang-Landau (WL) simulation [33] in the space of minimal energies, where $\varepsilon_{\min}$ plays the role of energy of a given configuration of A. (ii) Next, one performs a random walk in the space of configurations of A with weights

$$w(A) = 1/\tilde{\rho}[\varepsilon_{\min}(A)] \quad (2)$$

to sample $Z(\beta)$ for all β via the estimator $Z(\beta|A)/w(A)$. This weight function not only allows visits to all classes of configurations, but allows each class to be visited approximately the same number of times, yielding a flat histogram.

The above principle can be viewed as a minimal necessary requirement for sampling energy space. It yields a coarse-grained view of energy space disregarding any differentiation between configurations belonging to a given class. Below we show that the application of this principle leads to remarkably accurate results.

Testing the algorithm. The convergence properties of the WL algorithm that constitutes the first stage of our algorithm have been widely studied. We note that the random walk in the second stage of our algorithm (with fixed auxiliary DOS) fulfills, *by fiat*, detailed balance. We test the convergence of our algorithm on the (a) Ising and (b) Potts models. Finally, we present continuous temperature results for the Falicov-Kimball model as a prominent example of FCCF.

The Ising model with nearest-neighbor interactions on a square lattice is an ideal benchmark for testing new algorithms. This model is biseparable, which allows direct application of our algorithm. For a given configuration of spins on one sublattice (subsystem A) the conditional partition function corresponding to all spin configurations on the second sublattice (subsystem B) is simply $Z(\beta|A) = 2^{N/2}[\cosh(2\beta J)]^{N_2}[\cosh(4\beta J)]^{N_4}$, where J is the coupling between spins, $N/2$ is the number of sites on sublattice B, and N_2 and N_4 are the number of spins on sublattice B subject to the nearest-neighbor fields with absolute values $2J$ and $4J$, respectively, for a given configuration on sublattice B. Figure 1 shows the continuous temperature dependence of the specific

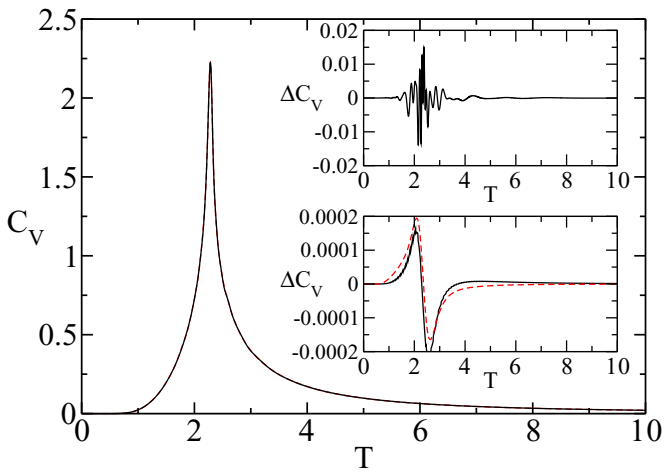

FIG. 1. Specific heat C_V per site for the 2D Ising model using our algorithm (black) and the WL (red dashed) algorithm. The lattice size is 60×60 with PBCs. The upper inset shows the difference between these lines. The lower inset shows the averaged error in our sampling (black line) for an 8×8 system with PBCs. The red dashed line depicts the error generated by temperature shifting exact data (see the text).

heat, obtained with our algorithm and the WL algorithm for a 60×60 site lattice [with periodic boundary conditions (PBCs)]. The two curves show good agreement, indicating the sufficiency of our prescribed importance sampling principle.

We note that, in principle, errors (top inset, Fig. 1) may be generated during both simulation stages in our algorithm. The properties of the WL algorithm have been been well studied in the literature (see Refs. [39–42]). To quantify errors arising during our importance sampling and to separate these from errors due to the first (i.e., WL) stage of our algorithm we study a 8×8 system for which the exact auxiliary DOS can be directly obtained by summing the 2^{32} sublattice A configurations. For the analysis, we performed 1600 independent runs of importance sampling (second stage) with this DOS, each consisting of 10^6 MC steps per spin. The evaluated C_V were compared with *exact* results for the system obtained using Beale's procedure [43]. The resulting average error is presented in the lower inset of Fig. 1 (black curve). We note first that the envelope of the error is analogous to that obtained for the 60×60 site system (top inset in Fig. 1). The persistence of the error for the large number of runs reveals a very small systematic error arising in the sampling scheme. We identify this to stem primarily from a simple overestimation of the finite-size critical temperature T_C^*. To see this, we determine the temperatures corresponding to the maxima of C_V curves for all runs and find them to follow a Gaussian distribution around an average value of $T_C^* = 2.362\,08$, while the exact value is $T_{C,\text{exact}}^* = 2.361\,98$. Shifting the exact data by the obtained difference $\Delta T_C^* = 0.000\,11$ produces an error depicted by the red dashed curve in the lower inset of Fig. 1, in striking agreement with that obtained from the Monte Carlo runs. Interestingly, a recent error analysis of the standard WL algorithm [44] showed the same type of small systematic shift for the discussed 8×8 Ising system, with the averaged temperature shift of C_V peak locations being of the same order ($\Delta T_C^* = 0.000\,43$). We emphasize here that the systematic errors in both the WL scheme and our method are very small (of the same order 10^{-4} for the 8×8 system) for practical purposes. While ways to remedy the errors have been proposed for the WL algorithm in Ref. [44], we leave this issue for our algorithm open for future study.

We now highlight some important distinctions between the WL and our algorithm. A single step in the second stage of our algorithm entails accumulation of full thermodynamic information from all configurations of subsystem B (a remarkable 2^{1800} configurations in the 60×60 site simulation) during each update move on subsystem A. In contrast, the standard WL algorithm updates information from one configuration per move. However, this exponential update comes at the initial cost of first obtaining the auxiliary DOS via a WL procedure. Interestingly, the auxiliary DOS has to be determined with essentially a higher histogram flatness requirement than the system DOS in the direct WL simulation of the Ising model. This has two sources. First, the subsystem Hamiltonian with energies $\varepsilon_{\min}$ contains multispin and long-range interactions unlike the original Ising interaction Hamiltonian. Second is sensitivity to the precision of the auxiliary DOS. We have found that adding small fluctuations to $\tilde{\rho}(\epsilon_{\min})$ rapidly deteriorates results. Hence, our algorithm should not be viewed as an alternative to the standard WL algorithm for classical

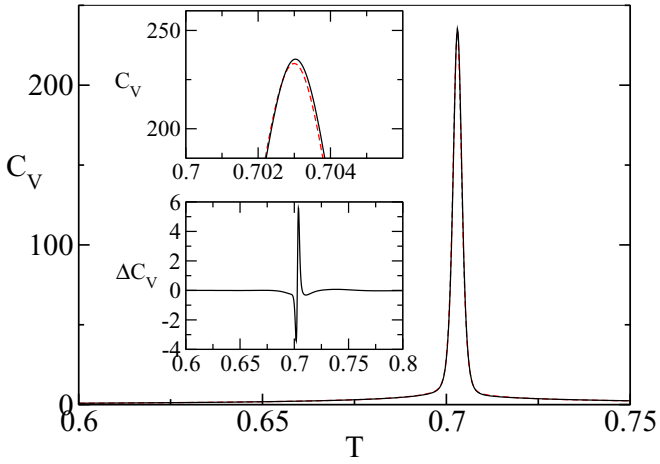

FIG. 2. Specific heat C_V per site for the 30×30 site 2D ten-state Potts model using our algorithm (black) and the WL (red dashed) algorithm. The lower inset shows the difference between these lines. The upper inset is a close-up of the transition region.

models. Finally, while the WL algorithm directly outputs the system DOS, our algorithm yields the partition function.

To illustrate that our algorithm successfully simulates discontinuous phase transitions, we consider the ten-state Potts model on the square lattice with a nearest-neighbor interaction. Biseparability here may be shown in a similar fashion as in the Ising model. However, the form of the conditional partition function is more complicated than for the Ising model due to the multitude of values of local fields set by configurations of nearest-neighbor Potts spins. A comparison in Fig. 2 of the specific heat, obtained by the WL algorithm and our algorithm, for a 30×30 site lattice with PBCs reveals the effectiveness of our sampling scheme in this case as well. There is a small shift in T_C^* between the WL algorithm and our method given by $\Delta T_C^* = 0.000\,04$, which in combination with the steepness of the peak at the transition contributes mainly to the errors shown in the upper inset in Fig. 2.

Now consider the Falicov-Kimball model (FKM)

$$H_{\text{FK}} = -t \sum_{\langle i,j \rangle} c_i^\dagger c_j + U \sum_i n_i^c n_i^d - \mu \sum_i n_i^c \qquad (3)$$

of FCCFs, where $c_i^\dagger$ ($d_i^\dagger$) is the creation operator of mobile (immobile) fermions, $n_i^c = c_i^\dagger c_i$, $n_i^d = d_i^\dagger d_i$, t is the nearest-neighbor hopping integral, U is the Hubbard on-site interaction, and μ is the chemical potential for c fermions (the number N_d of immobile fermions is fixed). For a given configuration of immobile fermions $\{n_i^d\}$, the set of single-particle eigenenergies $\{\varepsilon_l\}$ of mobile fermions is readily obtained, rendering the model biseparable.

Unlike for classical models above, an efficient standard WL algorithm is not directly applicable. We consider the FKM on a $N = L^2$ square lattice under PBCs with $L = 16$, at half filling ($\mu = U/2$ and $N_d = N/2$), for different values of U. Under these conditions the FKM undergoes, at low temperatures, a transition to the charge-density wave ordered state, with a $\mathbf{Q} = (\pi,\pi)$ ordering wave vector. The transition is continuous for large U, with some evidence of discontinuous transitions for small U [45,46]. Our algorithm yields all-temperature

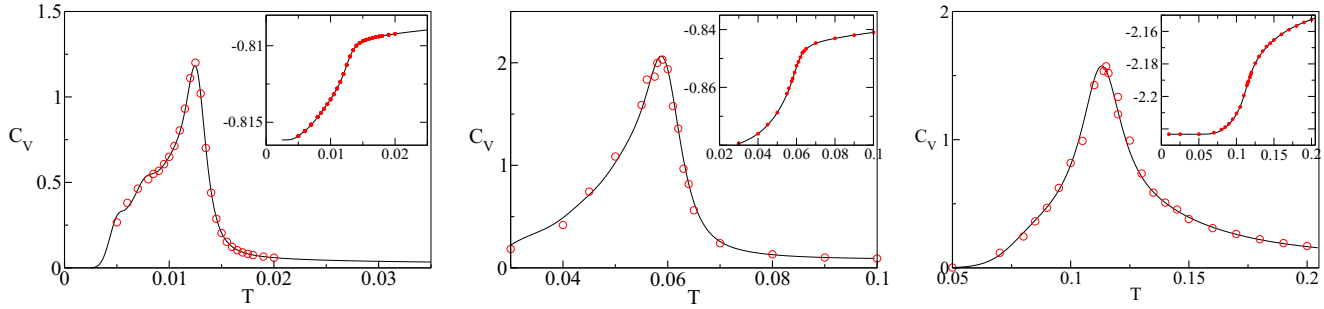

FIG. 3. Specific heat C_V per site for the 2D Falicov-Kimball model at half filling obtained with our algorithm (black line) and the Metropolis (red circles) algorithm. The electron-ion on-site repulsion U/t is equal to (from the left panel to the right) 0.25 (weak coupling), 1 (moderate coupling), and 8 (strong coupling). The lattice size is 16×16 with PBCs. The insets show a comparison of internal energies.

results unlike standard Metropolis sampling over the immobile fermions. Moreover, since the Metropolis algorithm suffers from slow kinetics at discontinuous phase transitions, our method should be particulary useful in further studies of the small-U regime. Here we restrict ourselves to proving that our sampling principle works. In Fig. 3 we compare the specific heat and internal energy obtained by Metropolis sampling and our algorithm. The results are within the accuracy of results of the local update Metropolis algorithm. We mention that, solely for generating the auxiliary DOS, we discretized minimal energies of the total system into energy bins of width $0.005t$ or $0.001t$ and checked for convergence.

Conclusion. We have presented a Monte Carlo algorithm for fermions coupled to classical fields based on Wang-Landau-like sampling in contradistinction to common Metropolis sampling. This allows, in principle, one to overcome drawbacks of Metropolis schemes. Importantly, it fully exploits all available thermodynamic information per diagonalization step, information that is mostly wasted in other MC schemes. Our scheme is based on the notion of minimal energy attainable for a given classical field configuration. We devised a minimal rule for importance sampling: that all such minimal energies be visited by the algorithm. We mention that while our sampling principle yields satisfactory results in the presented examples, more generally supplementary conditions may be required in other cases. However, our results show that, in principle, accumulation of conditional partition functions at all temperatures at once using simple temperature-independent importance sampling is possible. An intriguing possible application of our method is to quenched disorder problems, where free energies need to be efficiently averaged over disorder realizations.

Standard MC schemes update information from single or effectively at most M configurations (for an M-site system) per move (in the N-fold way algorithm [47]). Significantly, our algorithm updates information from an exponential number of configurations per move [$\sim \exp(\alpha M)$ for constant α]. Clearly, in terms of complexity, no algorithms may update more information. Importantly, the cost of this is standard for MC simulations: 10^6 (Ising) to 10^8 (Potts and FKM) steps per spin for importance sampling and a similar number for the weight function generation.

Acknowledgments. We thank Mathias Troyer for interesting and useful comments. Support from the (Polish) National Science Center Grant No. DEC-2011/03/B/ST2/01903, Spanish MINECO grant FOQUS (No. FIS2013-46768-P) and Severo Ochoa programme (SEV-2015-0522), Generalitat de Catalunya Grant No. (2014 SGR 874), and Fundació Cellex is acknowledged.

[1] C. Zener, Phys. Rev. **82**, 403 (1951).
[2] P. W. Anderson and H. Hasegawa, Phys. Rev. **100**, 675 (1955).
[3] P.-G. de Gennes, Phys. Rev. **118**, 141 (1960).
[4] J. Schliemann, J. Konig, and A. H. MacDonald, Phys. Rev. B **64**, 165201 (2001).
[5] M. M. Maśka, M. Mierzejewski, A. Ferraz, and E. A. Kochetov, J. Phys.: Condens. Matter **21**, 045703 (2009).
[6] T. Holstein, Ann. Phys. (N.Y.) **8**, 325 (1959).
[7] A. J. Millis, R. Mueller, and B. I. Shraiman, Phys. Rev. B **54**, 5389 (1996).
[8] J. K. Freericks, M. Jarrell, and G. D. Mahan, Phys. Rev. Lett. **77**, 4588 (1996).
[9] J. K. Freericks and V. Zlatić, Rev. Mod. Phys. **75**, 1333 (2003).
[10] J. Hubbard, Proc. R. Soc. London A **276**, 238 (1963); **281**, 401 (1964).
[11] L. M. Falicov and J. C. Kimball, Phys. Rev. Lett. **22**, 997 (1969).
[12] A. Kitaev, Ann. Phys. (N.Y.) **321**, 2 (2006).
[13] R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, Phys. Rev. D **24**, 2278 (1981).
[14] J. E. Hirsch and R. M. Fye, Phys. Rev. Lett. **56**, 2521 (1986).
[15] A. Mukherjee, N. D. Patel, S. Dong, S. Johnston, A. Moreo, and E. Dagotto, Phys. Rev. B **90**, 205133 (2014).
[16] S. Ulam, R. D. Richtmyer, and J. von Neumann, Statistical methods in neutron diffusion, Los Alamos Scientific Laboratory Report No. LAMS-551, 1947 (unpublished).
[17] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[18] E. Dagotto, *Nanoscale Phase Separation and Colossal Magnetoresistance: The Physics of Manganites and Related Compounds* (Springer, Berlin, 2003).
[19] Y. Motome and N. Furukawa, J. Phys. Soc. Jpn. **68**, 3853 (1999); **69**, 3785 (2000); **70**, 3186 (2001).
[20] Y. Motome and N. Furukawa, J. Phys. Soc. Jpn. **72**, 2126 (2003); **73**, 1482 (2004); G. Alvarez, C. Sen, N. Furukawa, Y. Motome, and E. Dagotto, Comput. Phys. Commun. **168**, 32 (2005).

[21] G. Alvarez, P. K. V. V. Nukala, and E. D'Azevedo, J. Stat. Mech. (2007) P08007.
[22] A. Weiße, Phys. Rev. Lett. **102**, 150604 (2009).
[23] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).
[24] U. Wolff, Phys. Rev. Lett. **60**, 1461 (1988).
[25] H. G. Evertz, G. Lana, and M. Marcu, Phys. Rev. Lett. **70**, 875 (1993).
[26] N. V. Prokof'ev, B. V. Svistunov, and I. S. Tupitsyn, Sov. Phys. JETP **87**, 310 (1998).
[27] A. W. Sandvik, Phys. Rev. B **59**, R14157 (1999); A. Dorneich and M. Troyer, Phys. Rev. E **64**, 066701 (2001).
[28] O. F. Syljuåsen and A. W. Sandvik, Phys. Rev. E **66**, 046701 (2002); O. F. Syljuåsen, *ibid.* **67**, 046701 (2003).
[29] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).
[30] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992); Phys. Lett. B **267**, 249 (1991).
[31] P. M. C. de Oliveira, T. J. P. Penna, and H. J. Herrmann, Braz. J. Phys. **26**, 677 (1996).
[32] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996); E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992); A. P. Lyubartsev *et al.*, J. Chem. Phys. **96**, 1776 (1992).
[33] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001); Phys. Rev. E **64**, 056101 (2001).
[34] M. Troyer, S. Wessel, and F. Alet, Phys. Rev. Lett. **90**, 120201 (2003).
[35] R. T. Scalettar, D. J. Scalapino, and R. L. Sugar, Phys. Rev. B **34**, 7911 (1986).
[36] S. Duane, A. Kennedy, B. J. Pendleton, and D. Roweth, Phys. Lett. B **195**, 216 (1987).
[37] J. L. Alonso, L. A. Fernández, F. Guinea, V. Laliena, and V. Martín-Mayor, Nucl. Phys. B **596**, 587 (2001).
[38] K. Barros and Y. Kato, Phys. Rev. B **88**, 235101 (2013).
[39] C. Zhou and R. N. Bhatt, Phys. Rev. E **72**, 025701(R) (2005).
[40] D. P. Landau, S.-H. Tsai, and M. Exler, Am. J. Phys. **72**, 1294 (2004).
[41] H. K. Lee, Y. Okabe, and D. P. Landau, Comput. Phys. Commun. **175**, 36 (2006).
[42] R. E. Belardinelli and V. D. Pereyra, J. Chem. Phys. **127**, 184105 (2007).
[43] P. D. Beale, Phys. Rev. Lett. **76**, 78 (1996).
[44] A. A. Caparica and A. G. Cunha-Netto, Phys. Rev. E **85**, 046702 (2012).
[45] M. M. Maśka and K. Czajka, Phys. Rev. B **74**, 035109 (2006).
[46] M. Žonda, P. Farkašovský, and H. Čenčariková, Solid State Commun. **149**, 45 (2009).
[47] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz, J. Comput. Phys. **17**, 10 (1975).