Rapid Communications

Clock Monte Carlo methods

Manon Michel,[1,2,*] Xiaojun Tan,[3] and Youjin Deng[3,4,†]

[1]*Centre de Mathématiques Appliquées, UMR 7641, École Polytechnique, Palaiseau, France*

[2]*Orange Labs, 44 Avenue de la République, CS 50010, 92326 Châtillon Cedex, France*

[3]*Hefei National Laboratory for Physical Sciences at Microscale and Department of Modern Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*

[4]*CAS Center for Excellence and Synergetic Innovation Center in Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*

(Received 30 June 2017; revised manuscript received 22 March 2018; published 25 January 2019)

We propose the clock Monte Carlo technique for sampling each successive chain step in constant time. It is built on a recently proposed factorized transition filter and its core features include its O(1) computational complexity and its generality. We elaborate how it leads to the clock factorized Metropolis (clock FMet) method, and discuss its application in other update schemes. By grouping interaction terms into boxes of tunable sizes, we further formulate a variant of the clock FMet algorithm, with the limiting case of a single box reducing to the standard Metropolis method. A theoretical analysis shows that an overall acceleration of $O(N^{\kappa})$ $(0 \leqslant \kappa \leqslant 1)$ can be achieved compared to the Metropolis method, where N is the system size and the κ value depends on the nature of the energy extensivity. As a systematic test, we simulate long-range $O(n)$ spin models in a wide parameter regime: for $n = 1, 2, 3$, with disordered, algebraically decaying or oscillatory Ruderman-Kittel-Kasuya-Yosida-type interactions and with and without external fields, and in spatial dimensions from $d = 1, 2, 3$ to the mean field. The O(1) computational complexity is demonstrated, and the expected acceleration is confirmed. Its flexibility and its independence from the interaction range guarantee that the clock method would find decisive applications in systems with many interaction terms.

DOI: 10.1103/PhysRevE.99.010105

Markov-chain Monte Carlo methods (MCMC) are powerful tools in many branches of science and engineering [1–8]. For instance, MCMC plays a crucial role in the recent success of AlphaGo [9], and appears as a keystone of the potential next deep learning revolution [10,11]. To estimate high-dimensional integrals, MCMC generates a chain of random configurations, called samples. The stationary distribution is typically a Boltzmann distribution and the successive moves depend on the induced energy changes. Despite a now long history, the most successful and influential MCMC algorithm remains the founding Metropolis algorithm [12] for its generality and ease of use, ranked as one of the top ten algorithms in the 20th century [13].

The Metropolis algorithm has, however, two major limitations. First, nearby samples can be highly correlated and, around the phase transition, the simulation efficiency drops quickly as the system size N increases. Second, an attempted move requires calculating the induced total energy change, leading to expensive computational complexities of up to $O(N)$ for systems with long-range interactions. This issue is also very acute in machine learning, where likelihood evaluations [14] scale with the number of data points.

Enormous effort has been devoted to circumventing the two limitations. Various efficient update schemes have been designed, including the celebrated cluster and worm algorithms [15–17], and the event-chain (EC) irreversible method [18,19]. Several techniques are also available in reducing the computational complexity for specific algorithms and systems. An "early-rejection" scheme was mentioned [2], which is nevertheless of $O(N)$ complexity. Making use of the particular feature that each bond is treated independently in the cluster-update scheme [20], Luijten and Blöte [21] applied an efficient sampling procedure to place occupied bonds, instead of visiting each bond sequentially and throwing a random number to decide its status. The Luijten-Blöte cluster algorithm has O(1) complexity [21–23], and has been generalized to quantum systems [24]. Recently, an EC algorithm was proposed for long-range soft-sphere systems [25].

In this Rapid Communication, we propose a general "clock" MC method, which has a constant-time sampling and can be applied to various update schemes. The core ingredient is the factorized Metropolis filter proposed in Ref. [19]. In particular, an algorithm of O(1) computational complexity is formulated in the framework of the local and most general update scheme, which we call the clock factorized Metropolis (FMet) algorithm. By grouping the interaction terms into boxes of tunable sizes, we further obtain a variant of the clock FMet algorithm for efficiency optimization. The limiting case of a single box recovers the Metropolis method, directly illustrating the generality of the clock FMet algorithm. We also discuss how the clock technique acts as a common ground for existing exact complexity reduction methods and present in particular its implementation in the EC update scheme.

*manon.michel@normalesup.org

†yjdeng@ustc.edu.cn

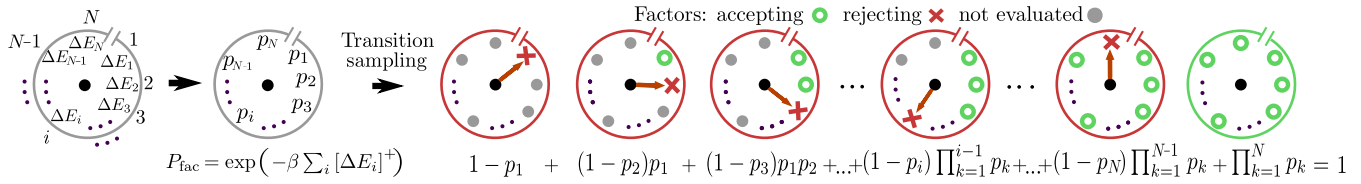

FIG. 1. Sketch of the clock sampling technique. The inhomogeneous Bernoulli process described by Eq. (2) can be decomposed into $N+1$ *clocks*, where the ith *clock* ($i \leqslant N$) *alarm* of the first occurrence of a rejection event at time i and the $(N+1)$th *clock*, with no alarm, represents an acceptance event. Such a process can be sampled within O(1) computational complexity.

While gaining an O(N) speeding-up in computational effort, the clock FMet algorithm can suffer from a lower acceptance probability than the Metropolis method. The *overall* acceleration comes from the compromise of these two effects. We provide a systematic performance analysis by classifying the system into three types of extensivity: *strict extensivity, marginal extensivity, and subextensivity*, and show that an overall acceleration can be achieved up to O(N) for the strict extensivity and O(N^κ) ($1 > \kappa \geqslant 0$) for the other two. As other exact complexity reduction methods belong to the algorithmic clock class, this analysis also applies to these techniques. Finally, we extensively simulate long-range O(n)-spin models in a wide parameter regime: for $n = 1, 2, 3$, with disordered, algebraically decaying or oscillatory Ruderman-Kittel-Kasuya-Yosida (RKKY)-type interactions, with and without external fields, and in spatial dimensions from $d = 1, 2, 3$ to the mean field. The O(1) computational complexity is demonstrated, and the expected acceleration is confirmed. These achievements are based on the complementary combination of the factorized Metropolis filter, O(1) sampling procedures, and the grouping trick.

I. CLOCK FMET ALGORITHM

Consider a system described by a collection of states $\mathcal{S}$ with a Boltzmann weight $\pi(\mathcal{S}) \propto \exp[-\beta E(\mathcal{S})]$, with $\beta = 1/k_{\rm b}T$ the inverse temperature. The energy $E(\mathcal{S}) = \sum_i E_i(\mathcal{S})$ is the sum of all interaction terms that are pairwise or more generally in many-body groups. At each step, the Metropolis algorithm attempts to update a state $\mathcal{S}$ into another $\mathcal{S}'$ with acceptance probability

$$P_{\rm Met} = \min\left(1, \frac{\pi(\mathcal{S}')}{\pi(\mathcal{S})}\right) = \exp(-\beta[\Delta E_{\rm tot}]^+), \qquad (1)$$

with $[x]^+ = \max(0, x)$. Evaluating the induced energy change $\Delta E_{\rm tot} \equiv \sum_i \Delta E_i$ requires a costly computation of all the involved interactions. Therefore, we focus now on the factorized Metropolis filter [19]

$$P_{\rm fac} = \prod_i p_i(\mathcal{S} \to \mathcal{S}') = \prod_i \exp(-\beta[\Delta E_i]^+), \qquad (2)$$

which also satisfies the detailed-balance condition $\pi(\mathcal{S})p(\mathcal{S} \to \mathcal{S}') = \pi(\mathcal{S}')p(\mathcal{S}' \to \mathcal{S})$. Hereinafter we omit the dependence on $\mathcal{S} \to \mathcal{S}'$ except in case it hinders the clarity. The factorized filter is a key component of the recent EC methods, as it allows one to extract interesting system symmetries. On a more general level, the factorization of transition rates can also play an important physical role in dynamical studies [26]. A crucial feature of Eq. (2) is the consensus rule: As the transition probability $P_{\rm fac}$ is a product of independent factors p_i, an attempted move is accepted *only if* all the factors give permission (Fig. 1). This leads to a lower acceptance probability in Eq. (2) than in Eq. (1). However, we show here how it plays a key role in designing the clock technique that dramatically reduces the computational complexity from O(N) to O(1), greatly improving the *overall* performance.

Without loss of generality, we illustrate the clock FMet method in the example of a long-range O(n) model of $N+1$ spins, with the Hamiltonian

$$\mathcal{H} = -c(N) \sum_{i<j} J_{ij} \boldsymbol{S}_i \cdot \boldsymbol{S}_j \quad (|\boldsymbol{S}| = 1), \qquad (3)$$

with $\boldsymbol{S}$ unit vectors in $\mathbb{R}^n$. For $n = 1, 2, 3$, one has the Ising, XY, and Heisenberg models, respectively. The coupling strength J_{ij} depends on distance r_{ij}, and can be ferromagnetic ($J_{ij} > 0$), antiferromagnetic ($J_{ij} < 0$), or disordered. There are in total $N(N+1)/2$ interaction terms. The normalization constant $c(N)$, scaling typically in $1/N^\alpha$ ($1 \geqslant \alpha \geqslant 0$), is to ensure the energy extensivity, which, e.g., is $1/N$ for a mean-field ferromagnet but $1/\sqrt{N}$ for the Sherrington-Kirkpatrick model [27,28]. An attempted move is to flip or rotate a randomly chosen spin $\boldsymbol{S}_j$. This leads to an energy change $-c(N)\sum_{i\neq j} J_{ij}(\boldsymbol{S}'_j - \boldsymbol{S}_j) \cdot \boldsymbol{S}_i$, which requests an O($N$) computation in the Metropolis algorithm.

A straightforward implementation of Eq. (2) is as follows. One orders the factor terms from $i = 1$, sequentially samples the rejection of each factor i with probability $1 - p_i$, and stops at the first-rejecting factor $i_{\rm rej}$; if no rejection is sampled until factor N, the move is accepted. This is analogous to an inhomogeneous Bernoulli process of rate p_i, as illustrated in Fig. 1, where the ith clock, with probability $P_{\rm rej}(i) = (1 - p_i)\prod_{k=1}^{i-1} p_k$, represents the event for the ith factor to be first-rejecting. Instead of sequentially sampling each factor, one can also evaluate cumulative probability $F_k = \sum_{k'=1}^{k} P_{\rm rej}(k')$ and directly obtain the $i_{\rm rej} = i$ value by solving $F_{i-1} < \nu \leqslant F_i$ with a single random number $\nu \in (0, 1]$. Nevertheless, the individual probabilities p_k depend *a priori* on a local configuration $(\boldsymbol{S}_i, \boldsymbol{S}_j)$, and each move still requires an average number $\mathcal{C} \sim$ O(N) of p_k evaluations.

To avoid these costly evaluations, we introduce a bound Bernoulli process with a *configuration-independent* probability $\widehat{p}_i$, so that $1 - \widehat{p}_i \geqslant 1 - p_i(\mathcal{S} \to \mathcal{S}')$, as done in Ref. [21]. An actual rejection at a factor i corresponds to a bound

rejection once resampled with relative probability

$$p_{i,\mathrm{rel}} = (1 - p_i)/(1 - \widehat{p}_i). \tag{4}$$

At each factor i, three events are possible: (A_1) bound acceptance with $p_i^{A_1} = \hat{p}_i$, (A_2) bound rejection and resampling rejection with $p_i^{A_2} = (1 - \hat{p}_i)(1 - p_{i,\mathrm{rel}})$, and (R) bound rejection and resampling acceptance with $p_i^{R} = (1 - \hat{p}_i)p_{i,\mathrm{rel}}$. Sampling the ith clock, i.e., the first-rejection at factor i, is then replaced by sampling a random path of events (A_1) or (A_2) for $k \leqslant i - 1$ and a first event (R) at i, as described by

$$P_{\mathrm{rej}}(i) = p_i^{R} \prod_{k=1}^{i-1} \left(p_k^{A_1} + p_k^{A_2}\right). \tag{5}$$

As the bound $\hat{p}_i$'s are configuration independent, the bound cumulative probabilities $\widehat{F}_i$ can be analytically calculated or tabulated. Initializing $i_{\widehat{\mathrm{rej}}} = 0$, the next bound rejection $i_{\widehat{\mathrm{rej}}}$ is updated to i by solving

$$\widehat{F}_{i-1} < \nu\left(1 - \widehat{F}_{i_{\widehat{\mathrm{rej}}}}\right) \leqslant \widehat{F}_i, \tag{6}$$

and the resampling is then applied. This is done within an O(1) complexity. If no actual rejection occurs, i.e., event (A_2), the procedure is repeated until an event (R) (actual rejection) is sampled or until $\nu(1 - \widehat{F}_{i_{\widehat{\mathrm{rej}}}}) > \widehat{F}_N$ (actual acceptance). The overall complexity $\mathcal{C}$ identifies now with the average number of attempted bound rejections $\sim \mathrm{O}(\ln P_B / \ln P_{\mathrm{Fac}}) \sim \mathrm{O}(1)$ if the bound consensus probability $P_B = \prod \widehat{p}_i$ scales with N as P_{Fac}. For a homogeneous case $\widehat{p}_i \equiv \widehat{p}$, Eq. (6) reduces to $i = i_{\widehat{\mathrm{rej}}} + \lfloor 1 + \ln(\nu)/\ln(\widehat{p}) \rfloor$, which can be easily adapted to inhomogeneous bound probabilities by ordering the factors decreasingly with $\widehat{p}_i$ and by replacing $\widehat{p}$ by $\widehat{p}_{i_{\widehat{\mathrm{rej}}}+1}$ [29]. Alternatively, one can directly generate the whole list of bound rejection events by the Walker method [30,31] or its Fukui-Todo extension [22], and then sequentially apply the resampling.

Algorithm 1 summarizes a clock FMet method for a long-range spin system. For Hamiltonian (3), $\widehat{p}_k$ can be taken as a function of distance r_{ij} as $\widehat{p}(r_{ij}) = \exp(-2\beta c(N)|J_{ij}|)$.

The clock method can be applied to any transition probability expressed as a product of independent factors, as the one proposed in [26] for instance. The factorized Metropolis filter, in addition to a maximal acceptance rate factorwise, presents the following advantage. As each factor in Eq. (2) can contain an *arbitrary* number of interactions, we introduce a variant of clock FMet algorithm in which the interactions are grouped into *boxes* b of *tunable* sizes B_b as

$$P_{\mathrm{fac}}^{\mathrm{Box}} = \prod_b \exp\left(-\beta \left[\sum_{i=1}^{B_b} \Delta E_{b_i}\right]^+\right). \tag{7}$$

It leads to new optimization possibilities (e.g., how to group the interaction terms). If all the interactions are in a single box, one recovers the standard Metropolis method.

The clock technique has two important ingredients: the consensus rule and the resampling. Both the ingredients are *general*: They do not depend either on any specific configurations, or on factor ordering, or on energy functions, or on update schemes. For systems in a continuous volume, as soft spheres, one can introduce a grid [25] to which the clock technique is applied.

Algorithm 1 Clock factorized Metropolis (Clock FMet)

Draw a random spin j and a random move $\boldsymbol{S}_j \to \boldsymbol{S}'_j$
$i_{\widehat{\mathrm{rej}}} \leftarrow 0$ ▷ *Sample bound rejections starting from* $i_{\widehat{\mathrm{rej}}}$
while True **do**
  $i_{\widehat{\mathrm{rej}}} \leftarrow i$ ▷ *Next bound rejection* i *given by Eq. (6)*
  if $i_{\widehat{\mathrm{rej}}} > N$ **then**
    $\boldsymbol{S}_j \leftarrow \boldsymbol{S}'_j$ ▷ *Move accepted*
    Break
  else ▷ *Decide whether it is an actual rejection*
    $p_{i_{\widehat{\mathrm{rej}}},\mathrm{rel}} \leftarrow$ Eq.(4)
    if ran(0, 1) $\leqslant p_{i_{\widehat{\mathrm{rej}}},\mathrm{rel}}$ then
      Break ▷ *Move rejected*

II. GENERALIZATION TO OTHER UPDATE SCHEMES

We illustrate the generality of the clock method by discussing its application in the EC method for the O(n) spin model with $n \geqslant 2$ [32,33]. With an auxiliary *lifting* variable j that specifies the moving spin, the EC method proposes to rotate infinitesimally its angle as $\phi_j \to \phi_j + \mathrm{d}\phi$. Such a move is rejected by *at most* one spin i, owing to a continuous derivation of Eq. (2). This yields $p_i \to 1 - \lambda_i d\phi$ and $P_{\mathrm{rej}}(i) \to \lambda_i d\phi$. For each factor i, the rejection event, with distance $\delta_i\phi$, is thus ruled by a Poisson process (PP) of rate λ_i, a continuous derivation of the standard Bernoulli process. The spin j is then rotated by the minimum distance $\delta\phi = \min(\delta_i\phi)$, and the associated factor becomes the moving spin, i.e., $j \to i_{\min}$. For long-range interactions, evaluating $\delta_i\phi$ for all the factors becomes costly.

To derive the clock method, we introduce a bound Poisson process of total rate $\bar{\lambda} = \sum \bar{\lambda}_i$ ($\bar{\lambda}_i \geqslant \lambda_i$), evaluate a random bound rotation $\delta\widehat{\phi} = -\ln \nu/\bar{\lambda}$, sample the rejecting bound factor $i_{\widehat{\min}} = i$ with probability $\bar{\lambda}_i/\bar{\lambda}$, and resample it as an actual lift j to i according to $\lambda_i/\bar{\lambda}_i$ [Eq. (4)]. This comes down to the thinning method [34], already applied for soft-sphere systems [25] and for logistic regression in machine learning [35].

We also note that the cluster methods [15,16] factorize each interaction term independently as in Eq. (2). The resampling procedures in the extended cluster algorithms for long-range interactions and for quantum spin systems [21–24] can be understood as specific cases of the clock method.

III. PERFORMANCE ANALYSIS

We expect and numerically confirm in Fig. 2 that the standard Metropolis and the clock FMet algorithms have the same physical dynamics. The overall acceleration $\mathcal{A}$ in the latter comes then from the speeding-up in the complexity $\mathcal{C}$, corrected by the slowing-down γ due to a lower acceptance in the factorized filter (2), leading to $\mathcal{A} \sim \mathrm{O}(N/\mathcal{C}\gamma)$. Both effects can be characterized by the scaling of $\sum_i \max|\Delta E_i|$ and $\sum_i |\Delta E_i|$, as $\gamma = P_{\mathrm{Met}}/P_{\mathrm{Fac}}$ can be written as

$$\ln \gamma = \frac{\beta}{2}\left(\sum_i |\Delta E_i| - \left|\sum_i \Delta E_i\right|\right), \tag{8}$$

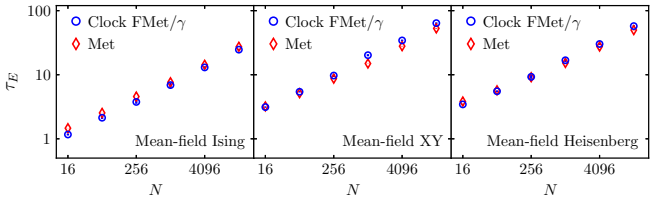

FIG. 2. Integrated autocorrelation time τ_E of the energy for the ferromagnetic mean-field $O(n)$ model expressed in units of system sweeps. The values for the clock FMet algorithm are normalized by γ [Eq. (8)].

and as $\mathcal{C} \sim \ln P_B / \ln P_{\text{Fac}} \sim \sum_i \max|\Delta E_i| / \sum_i |\Delta E_i|$. Depending on the nature of the energy extensivity and phase of the system, the sum $\sum_i |\Delta E_i|$ may diverge as size $N \to \infty$, while the sum $|\sum_i \Delta E_i|$ converges to a constant. This normally occurs in disordered systems with slowly decaying interactions, in which the "satisfied" and "unsatisfied" interaction terms compensate each other. The divergence of γ can be controlled by introducing enough compensation through boxes. For a constant size B, it increases the complexity to $\mathrm{O}(B|\ln P_B|)$, but leads to an acceleration $\sim \mathrm{O}(N/(B|\ln P_B|))$. By definition, $B \propto N$ would ensure a maximal energy compensation but an O(1) acceleration.

We classify the system into the three types of extensivity: *strict extensivity*, *marginal extensivity*, and *subextensivity*, for which $\sum_i \max|\Delta E_i|$ respectively scales as O(1), $\mathrm{O}(\ln N)$, and $\mathrm{O}(N^\alpha)$ $(1 > \alpha > 0)$. It is noted that for the subextensivity type, the total energy of the system remains extensive because of the sign cancellation of individual energy terms. We demonstrate that the clock FMet method of *tunable* constant box sizes B might achieve an overall acceleration:

(1) $\mathcal{A} \sim \mathrm{O}(N)$ for strict extensivity, directly from $\gamma \sim \mathrm{O}(1)$ and $\mathcal{C} \sim \mathrm{O}(1)$.

(2) $\mathcal{A} \sim \mathrm{O}(N^\kappa)$ $(0 \leqslant \kappa < 1)$ for subextensivity. Depending on the phase, $\ln\gamma$ may diverge, up to N^α. For the spin glass of algebraically decaying interaction as $1/r^\sigma$ $(\sigma < 1)$, we find that a box size $B \propto N/N^\omega$, with a fine-tuning exponent $0 < \omega < 1$, gives a sufficient compensation and an $\mathrm{O}(N^\kappa)$ $(\kappa \sim \omega - \alpha)$ acceleration.

(3) $\mathrm{O}(N/(\ln N)^2) < \mathcal{A}_{\text{margin}} < \mathrm{O}(N/\ln N)$ for marginal extensivity. We observe that setting B up to $\ln N$ can be necessary to control γ.

For frustrated systems, irrespective of which class they belong to, efficient cluster algorithms are normally unavailable due to the huge cluster sizes. Given the substantial acceleration for all the three classes of extensivity, the application of the clock FMet method is very promising.

IV. SIMULATIONS

We simulate three typical systems in statistical physics, including long-range Ising spin glass, the disordered O(n) model with random external fields, and the O(n) model with RKKY-type interactions. We record the number of energy evaluations $\mathcal{C}$ for each MC step, which for the Metropolis method is simply $\mathcal{C} = N$. We measure the integrated correlation times τ for magnetic susceptibility χ in the units of energy evaluations, and compute the overall acceleration $\mathcal{A}$ as the inverse ratio $\mathcal{A} = \tau_{\text{other}}/\tau_{\text{FMet}}$, where τ_{other} is for the Metropolis or the Luijten-Blöte (LB) cluster method. For the Metropolis method, it comes down to $\mathcal{A} = N/(\gamma\mathcal{C})$.

A. Long-range Ising spin glass

We consider a periodic one-dimensional (1D) spin glass defined by Eq. (3). The interactions decay algebraically as $J_{ij} = s_{ij}/r_{ij}^\sigma$ $(\sigma > 0)$, with $s_{ij} = \pm 1$ from a bimodal distribution. The normalization $c(N)$ is given by $c(N)^{-2} = \sum_{j>1} \langle J_{1j}^2 \rangle$. This system, with a tunable exponent σ, is particularly useful in revealing the crossover behavior from the low-dimensional to the mean-field spin glass [36–38]. For simplicity, the simulation is made at the mean-field critical temperature $\beta = 1$ [27]. Depending on the value of σ, we recover the three extensivity regimes, i.e., strict extensivity $(\sigma > 1)$, marginal extensivity $(\sigma = 1)$, and subextensivity $(\sigma < 1)$. We group the interaction terms following respective r_{ij} values and set the box size as $B = 2$ $(\sigma > 1)$, $\ln N$ $(\sigma = 1)$, and $N^{2(1-\sigma)}$ $(\sigma < 1)$. The results are shown in Fig. 3. For $\sigma > 1$, the computational complexity $\mathcal{C}$ converges to a constant, and a dramatic overall acceleration $\mathcal{A} \sim \mathrm{O}(N)$ is achieved. For $\sigma = 1$, we have $\mathcal{C} \sim \ln N^2$ and a significant acceleration $\mathcal{A}$ converging to $N/(\ln N)^2$. For $\sigma < 1$, both $\mathcal{C}$ and $\mathcal{A}$ increase sublinearly as N. The acceleration $\mathcal{A}$ drops as σ becomes smaller. Nevertheless, given the simplicity of the clock FMet method, the gained improvement is still significant. We also

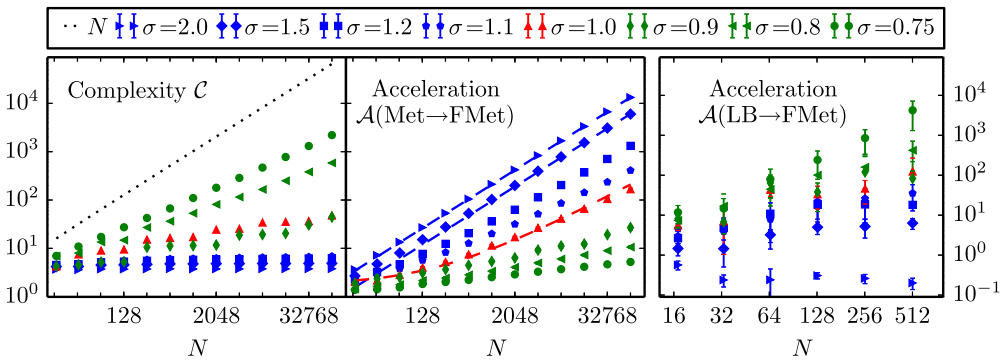

FIG. 3. Complexity $\mathcal{C}$ (left) and acceleration $\mathcal{A}$ for the clock FMet algorithm for the 1D $(\beta = 1)$ long-range Ising spin glass, compared to the Metropolis algorithm (middle) and the LB algorithm (right). The dashed blue and red lines respectively represent fits to aN and $aN/(\log N)^2 + b$.

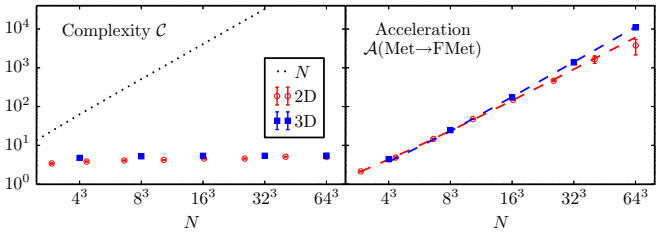

FIG. 4. Complexity $\mathcal{C}$ (left) and acceleration $\mathcal{A}$ (right) for the 2D ($\beta = 1$, red circle) and 3D ($\beta = 0.693$, blue square) RKKY Heisenberg spin systems, comparing the clock FMet method to the Metropolis algorithm. The dashed red and blue lines respectively represent fits to aN and $aN/\log N + b$.

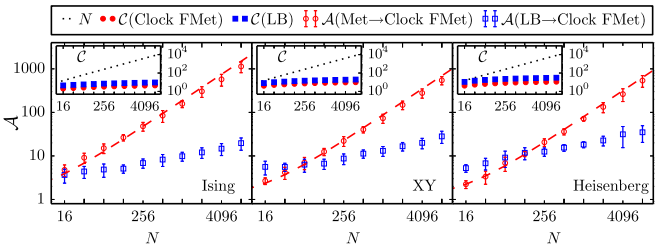

FIG. 5. Acceleration $\mathcal{A}$ for the disordered mean-field O(n) model in random fields at $\beta = n$ (complexity $\mathcal{C}$ in insets). The clock FMet method exhibits important acceleration $\mathcal{A}$, compared to both the Metropolis (red circles) and the LB cluster algorithm (blue squares). The dashed red line represents a fit to $aN + b$.

compare its performances with the LB algorithm, which confirms the superiority of the local Clock FMet for disordered systems. These results are fully consistent with the performance analysis.

B. RKKY-type interactions

We then consider the 2D and 3D Heisenberg models with oscillatory Ruderman-Kittel-Kasuya-Yoshida (RKKY) interactions $J_{ij} = J_0[\cos(2k_F r_{ij})/r_{ij}^d]\exp(-r_{ij}/\lambda)$, where d is the spatial dimension, k_F is the Fermi vector ($k_F \approx 4.91$ for the spin-glass system $CuMn$), and λ is the characteristic length in the damping term [37,39–43]. Due to their approximate description of real materials, rich behaviors, and important roles in bridging the experimental study of glassy materials and the spin-glass theory of short-range interactions [37,39], these systems are under extensive studies. For simplicity, we set $J_0 = 1$ and $k_F = \pi$, and take $\lambda = 3$ for 3D and $\lambda = \infty$ for 2D, so that the system is in the class of strict (3D) and marginal (2D) extensivities. The simulations are at $\beta(2D) = 1$ and $\beta(3D) = 0.693$, close to the critical temperature $\beta_c = 0.693\,003(2)$ for the 3D pure Heisenberg model [44]. Box sizes are set to 1 and the achieved acceleration is again $\mathcal{A} \sim O(N)$ for the strict extensivity, and $\mathcal{A} \sim O(N/\ln N)$ for the marginal extensivity, as illustrated in Fig. 4.

C. Disordered random-field model

Finally, we study a disordered mean-field O(n) model in a random external field. The interactions are partly disordered, i.e., $J_{ij} = 1$ for 90% of interactions while the remaining J_{ij} are drawn from a normal distribution with $\langle J_{ij}\rangle = 0$ and $\langle J_{ij}^2\rangle = 1$. A quenched random field is applied to each lattice site as $-\boldsymbol{h_i}\cdot\boldsymbol{S_i}$, where $\boldsymbol{h_i}$ is drawn from an n-dimensional normal distribution. The normalization is $c(N) = 1/N$, and the system belongs to the class of strict extensivity. Random-field models have applications in a wide range of physics [45–49], including the pinning of vortices in superconductors, Coulomb glass, the metal-insulator transition, and hysteresis and avalanche physics. In general spatial dimensions, the thermodynamic properties and phase transitions are still debated [50,51]. We perform simulations at the mean-field critical temperature $\beta = n$, with box sizes set to 1. The results are shown in Fig. 5. The clock FMet method clearly displays an O(N) acceleration over the Metropolis algorithm for all the Ising, XY, and Heisenberg models. It also exhibits some superiority ($\mathcal{A} \sim 50$ for large system sizes) compared to the LB cluster algorithm that already implements the clock technique and has an O(1) computational complexity. The central-limit theorem tells that, as temperature is lowered and/or the strength of the external fields is increased, the acceptance rate exponentially drops for clusters of large sizes in the LB algorithm, and thus this superiority would become more pronounced.

V. CONCLUSION

We introduce a general clock technique with O(1) computational complexity for each Monte Carlo step, and discuss its implementations in various update schemes, regrouping most existing complexity reduction techniques into a single algorithmic class. An important application is the clock FMet algorithm. This is made possible owing to the following three flexible features of the factorized filter (2). First, the consensus rule in Eq. (2) allows for the decision of the fate of a proposed move by an O(1) sampling procedure. Second, the equal generality of Eqs. (1) and (2) allows for a similar application range. Third, the factorization range flexibility given by the grouping trick in Eq. (7) allows for a semicontinuation from the Metropolis Eq. (1) to the factorized filter Eq. (2) and a control over the frustration present in the considered system.

The clock FMet algorithm and its variant with tunable box sizes can lead to significant or even dramatic acceleration $\mathcal{A}$. Depending on the system, theoretical analysis gives $\mathcal{A}$ up to $O(N)$, $O(N/\ln N)$, and $O(N^\kappa)$ $(1 > \kappa \geqslant 0)$ for respectively strict extensivity, marginal extensivity, and subextensivity. Moreover, as the Metropolis method can be understood as a limiting case of the clock FMet, the latter cannot be worse. This is confirmed by simulations of long-range O(n) models in a wide parameter range. Since these systems are under active studies and the simulations rely heavily on the Metropolis method, the clock FMet algorithm is readily available to explore their rich physics. From its simplicity and ease of use, we conclude that the clock technique is a serious candidate for tackling Monte Carlo scaling in all scientific fields.

ACKNOWLEDGMENTS

X.T. and Y.D. are thankful for support by the National Natural Science Foundation of China under Grant No. 11625522 and the Ministry of Science and Technology of China (under

Grant No. 2016YFA0301604). M.M. thanks the University of Science and Technology of China for its hospitality during which this work was initiated and partly done and is very grateful for the support from the Data Science Initiative, the Chaire BayeScale "P. Laffitte" and the PHC program Xu Guangqi (Grant No. 41291UF).

[1] D. M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995).
[2] D. Frenkel and B. Smit, *Understanding Molecular Simulation – From Algorithms to Applications* (Academic, San Diego, 1996).
[3] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, UK, 2000).
[4] T. Opplestrup, V. V. Bulatov, G. H. Gilmer, M. H. Kalos, and B. Sadigh, Phys. Rev. Lett. **97**, 230602 (2006).
[5] D. W. O. Rogers, Phys. Med. Biol. **51**, R287 (2006).
[6] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer Texts in Statistics (Springer, Berlin, 1999).
[7] J. S. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer Series in Statistics (Springer, Berlin, 2001).
[8] P. Glasserman, *Monte Carlo Methods in Financial Engineering* (Springer, New York, 2004).
[9] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot *et al.*, Nature (London) **529**, 484 (2016).
[10] R. M. Neal, *Bayesian Learning for Neural Networks* (Springer, Berlin, 1996).
[11] Z. Ghahramani, Nature (London) **521**, 452 (2015).
[12] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[13] J. Dongarra and S. Sullivan, Comput. Sci. Eng. **2**, 22 (2000).
[14] R. Bardenet, A. Doucet, and C. Holmes, JMLR, **18**, 1 (2017).
[15] R. H. Swendsen and J. S. Wang, Phys. Rev. Lett. **58**, 86 (1987).
[16] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989).
[17] N. Prokof'ev and B. Svistunov, Phys. Rev. Lett. **87**, 160601 (2001).
[18] E. P. Bernard, W. Krauth, and D. B. Wilson, Phys. Rev. E **80**, 056704 (2009).
[19] M. Michel, S. C. Kapfer, and W. Krauth, J. Chem. Phys. **140**, 054116 (2014).
[20] R. G. Edwards and A. D. Sokal, Phys. Rev. D **38**, 2009 (1988).
[21] E. Luijten and H. W. J. Blöte, Int. J. Mod. Phys. C **06**, 359 (1995).
[22] K. Fukui and S. Todo, J. Comput. Phys. **228**, 2629 (2008).
[23] E. Flores-Sola, M. Weigel, R. Kenna, and B. Berche, Eur. Phys. J. Spec. Top. **226**, 581 (2017).
[24] Y. Deng and H. W. J. Blöte, Phys. Rev. Lett. **88**, 190602 (2002).
[25] S. C. Kapfer and W. Krauth, Phys. Rev. E **94**, 031302(R) (2016).
[26] A. Hucht, Phys. Rev. E **80**, 061138 (2009).
[27] D. Sherrington and S. Kirkpatick, Phys. Rev. Lett. **35**, 1792 (1975).
[28] S. Kirkpatrick and D. Sherrington, Phys. Rev. B **17**, 4384 (1978).
[29] J. G. Shanthikumar, Eur. J. Oper. Res. **21**, 387 (1985).
[30] A. J. Walker, ACM Trans. Math. Software **3**, 253 (1977).
[31] G. Marsaglia, W. W. Tsang, and J. Wang, J. Stat. Software **11**, 1 (2004).
[32] M. Michel, J. Mayer, and W. Krauth, Europhys. Lett. **112**, 20003 (2015).
[33] Y. Nishikawa, M. Michel, W. Krauth, and K. Hukushima, Phys. Rev. E **92**, 063306 (2015).
[34] P. A. W. Lewis and G. S. Shedler, Naval Res. Logist. **26**, 403 (1979).
[35] A. Bouchard-Côté, S. J. Vollmer, and A. Doucet, J. Am. Stat. Assoc. **113**, 855 (2018).
[36] G. Kotliar, P. W. Anderson, and D. L. Stein, Phys. Rev. B **27**, 602(R) (1983).
[37] A. J. Bray, M. A. Moore, and A. P. Young, Phys. Rev. Lett. **56**, 2641 (1986).
[38] F. Beyer, M. Weigel, and M. A. Moore, Phys. Rev. B **86**, 014431 (2012).
[39] F. Matsubara and M. Iguchi, Phys. Rev. Lett. **68**, 3781 (1992).
[40] D. J. Priour Jr., E. H. Hwang, and S. Das Sarma, Phys. Rev. Lett. **95**, 037201 (2005).
[41] D. J. Priour, Jr. and S. Das Sarma, Phys. Rev. Lett. **97**, 127201 (2006).
[42] K. Szalowski and T. Balcerzak, Phys. Rev. B **77**, 115204 (2008).
[43] K. Kirkpatrick and T. Nawaz, J. Stat. Phys. **165**, 1114 (2016).
[44] Y. Deng, H. W. J. Blöte, and M. P. Nightingale, Phys. Rev. E **72**, 016128 (2005).
[45] A. I. Larkin, Zh. Eksp. Teor. Fiz. **58**, 1466 (1970) [Sov. Phys. JETP **31**, 784 (1970)].
[46] A. L. Efros and B. I. Shklovskii, J. Phys. C **8**, L49 (1975).
[47] T. R. Kirkpatrick and D. Belitz, Phys. Rev. Lett. **73**, 862 (1994).
[48] J. P. Sethna, K. Dahmen, S. Kartha, J. A. Krumhansl, B. W. Roberts, and J. D. Shore, Phys. Rev. Lett **70**, 3347 (1993).
[49] M. L. Rosinberg, G. Tarjus, and F. J. Perez-Reche, J. Stat. Mech. (2009) P03003.
[50] F. Krzakala, F. Ricci-Tersenghi, and L. Zdeborova, Phys. Rev. Lett. **104**, 207208 (2010).
[51] L. Leuzzi and G. Parisi, Phys. Rev. B **88**, 224204 (2013).