Multiparticle Moves in Acceptance Rate Optimized Monte Carlo

Tobias Neumann, Denis Danilov, and Wolfgang Wenzel*

Molecular Dynamics (MD) and Monte Carlo (MC) based simulation methods are widely used to investigate molecular and nanoscale structures and processes. While the investigation of systems in MD simulations is limited by very small time steps, MC methods are often stifled by low acceptance rates for moves that significantly perturb the system. In many Metropolis MC methods with hard potentials, the acceptance rate drops exponentially with the number of uncorrelated, simultaneously proposed moves. In this work, we discuss a multiparticle Acceptance Rate Optimized Monte Carlo approach (AROMoCa) to construct collective moves with near unit acceptance probability, while preserving detailed balance even for large step sizes. After an illustration of the protocol, we demonstrate that AROMoCa significantly accelerates MC simulations in four model systems in comparison to standard MC methods. AROMoCa can be applied to all MC simulations where a gradient of the potential is available and can help to significantly speed up molecular simulations. © 2015 Wiley Periodicals, Inc.

Introduction

Computer simulations are widely used to investigate a variety of physical and chemical processes on the atomic, molecular, or nanoscale.[1] For many problems, the interactions between the constituents of the system can be represented by classical force fields and molecular dynamics (MD) or Monte Carlo (MC) simulations can be used to simulate the time evolution of the system or its thermodynamic properties, respectively. Molecular simulations, which are often termed "virtual experiments," are routinely applied to many problems, including protein folding, morphology prediction of thin film or glassy materials.[2–11] With increasing computational power and continued efforts in the development of force fields and simulation protocols, molecular simulations have contributed significantly to the development of many scientific disciplines, ranging from astrophysics to biology and medicine.

Nevertheless, there remain limitations to the system size and time-scale of the problems that can be sensibly studied with readily accessible computational hardware using these methods. Development of special-purpose computers, such as ANTON,[12,13] have shown that efforts to improve the range of applicability of molecular simulation methods can offer insights into complex phenomena that are presently difficult to access experimentally. For these reasons, it remains both challenging and interesting to investigate the fundamental limitations of the present-day molecular simulation methods, to continue to broaden their applicability. The main limiting factor in MD simulations is the inherently small time step that arises as a result of the discretization of Newton's equation of motion. Despite much research in this field,[14,15] most generic MD simulations with atomic resolution the time steps of an individual integration step has been limited to the femtosecond-scale.

For many problems, MC simulations are an alternative approach to compute thermodynamic expectation values for the system. In the most widely used Metropolis MC approach,[16] a Markov chain is generated by repeatedly applying a random modification to the system, usually called a "move," which is either accepted or rejected according to an acceptance criterion that generally satisfies detailed balance. MC simulations thus do not suffer from short time steps—their limitation lies in the move-construction which severely limits MC-based protocols in many applications. In principle, arbitrary modifications of the system can be proposed in a single "move," but as the acceptance rate typically falls exponentially with increasing energy (compared with the present state of the system), the acceptance rate for arbitrary modification of complex systems is typically exponentially small. As a result, MC protocols are generally efficient for problems with smooth or discrete potentials or with weakly correlated degrees of freedom, but fail when collective effects play a role in the evolution of the system. It is generally possible to define "local" moves, which modify only a small fraction of the degrees of freedom of the system in a single step, but it is clear that such moves become inefficient for large system sizes.

Conversely, it is difficult to construct collective moves with high acceptance rates in condensed systems with thousands to hundreds of thousands of particles. A notable exception are specialized MC protocols, such as those proposed by Wang

T. Neumann, D. Danilov, W. Wenzel
Institute of Nanotechnology, Karlsruhe Institute of Technology, PO Box 3640, D 76021 Karlsruhe, Germany

Contract grant sponsor: Ministry of Science, Research and Arts of the State of Baden Württemberg, Germany; Contract grant sponsor: Universities of the State of Baden Württemberg, Germany

and Swendsen[17,18] for discrete spin systems, which construct large-scale moves with zero energy change. Unfortunately, such methods are not generally available for systems with continuous variables and complex potential functions (in particular with hard-core potentials). In the absence of such moves, the number of steps and computational effort required to generate an uncorrelated conformation, measured by the autocorrelation "time," increases rapidly with the system size. This is the equivalent of the short time step in MD, which results in a large number of energy/gradient evaluations to propagate the system as a whole along a relevant macroscopic reaction coordinate. In both MC and MD, the computational time is dominated by the energy/gradient evaluation, and it should be noted that for most classical potentials, the evaluation of the gradient is not significantly more expensive than the evaluation of the energy. As a result, the number of steps required to decorrelate a macroscopic variable is the measure of the computational efficiency of the method.

In MC methods, one might, therefore, attempt to reduce the autocorrelation time by combining many uncorrelated "local" moves to a "collective" move before evaluating the energy. Unfortunately, this straightforward approach fails, as the acceptance probability, which is then the product of the acceptance probabilities of the individual moves decreases rapidly if there is an admixture of a few energetically unfavorable moves in a such a "collective" move (again typical for hard-core potentials). Generalized Monte-Carlo algorithms, such as replica exchange,[19] simulated annealing (SA),[20] or multiple try Monte Carlo (MTM)[21,22] increase the efficiency for specific applications but do not overcome the fundamental bottleneck described above.

Since their introduction in the 1950s, there have been many propositions to increase the efficiency of MC simulations,[23] including force biased move construction for the simulation of water,[24,25] cluster MC algorithms,[26,27] and approaches using generalized ensembles and feedback effects.[28] The paradigmatic idea of many of these methods is exemplified by the famous Swendsen–Wang algorithm where collective moves with a zero energy change that affect a large fraction of the total system have been constructed. This approach requires an efficient calculation of the exact energy change of the move, which is possible only for discrete systems with short-range interactions.

To address a more general class of systems, a number of MC methods have been proposed where the choice of a move is biased by an energy estimator and all degrees of freedom are changed simultaneously.[24,25,29,30] The idea to use an energy estimator to construct efficient moves was originally proposed by Rao et al. in 1979[25] and later taken up in 1992 by Dereli[31] and Timonova et.al. for the modeling of diffusion and phase transitions.[29] As then the approach was also investigated by Ref. 30 and extended to associate a time-scale with the MC algorithm.[32,33]

In all of these methods, the construction of moves is biased toward decreasing energy. The strength of this temperature-dependent bias is chosen to reflect detailed balance, provided the estimate of the energy change of the proposed move is exact. As this not the case in general, the methods obey detailed balance only asymptotically in the limit of vanishing step size and the accuracy and efficiency of the method depends on the "critical" step size that can be chosen for a particular system. When the step size is chosen larger than the critical step size, these methods violate detailed balance and may fail to compute proper thermodynamic averages. If the critical step size below which the violation of detailed balance may be tolerated is small, the methods become computationally inefficient. Other advanced MC methods such as the cluster algorithms proposed by Liu and Luijten[34] or Krauth[35] circumvent this problem but are not applicable to molecular simulations.

Here, we investigate a generic MC protocol,[24,25,36] called Acceptance Rate Optimized Monte-Carlo (AROMoCa), which is applicable to continuous systems and avoids these problems. In AROMoCa, we construct moves composed of many degrees of freedom with acceptance rates close to unity but preserve detailed balance exactly. This manuscript is structured as follows: In the next section, we present a detailed derivation of the method, followed by the investigation of four representative systems of increasing complexity that demonstrate its effectiveness in comparison to generic MC methods and also compare AROMoCa to MD simulations. Finally, we discuss the relevance of the critical step size in previously proposed methods, such as force-bias Monte-Carlo (fbMC).[29,30] The aim of the MC methods discussed here is to describe the properties of systems in thermodynamic equilibrium. Other methods, such as kinetic MC,[37] which are used to describe nonequilibrium problems are not considered in this publication.

Methods

Detailed balance and biased move construction

As mentioned above, many MC algorithms are based on the construction of Markov chains describing the evolution of a fictitious system where each new element depends only on its predecessor: In each extension of the chain, a trial change to the system of interest (move) is proposed which is either accepted or rejected according to a method- and system-specific acceptance criterion. A sufficient but not necessary condition to compute proper thermodynamic expectation values is the detailed balance condition which postulates that: the move construction and acceptance has to be chosen in a way that the total rate of states moving from state q into state q', $\Gamma((q) \to (q'))$, equals the rate of the inverse move from q' to q. For a system following a Boltzmann distribution, which we will describe without loss of generality in the following, the transition rate from a state q to a state q' is the product of the transition probability, $W((q) \to (q'))$, and the probability to be in state q:

$$\Gamma((q) \to (q')) \quad W((q) \to (q')) \times \frac{1}{\mathcal{Z}} \exp(\quad \beta E(q)) \qquad (1)$$

where $E(q)$ is the potential energy of the state q, $\beta \quad 1/(k_B T)$ is the inverse temperature and $\mathcal{Z} \quad \sum_q \exp(\quad \beta E(q))$ is the partition function. The transition probability is the product of the probability to construct the move, $\pi((q) \to (q'))$ and the probability to accept the move, $\rho((q) \to (q'))$:

$$W((q) \to (q')) \quad \pi((q) \to (q'))\, \rho((q) \to (q')) \tag{2}$$

A sufficient condition to reach thermodynamic equilibrium, that is, $\Gamma((q) \to (q')) \quad \Gamma((q') \to (q))$ is the detailed balance criterion:

$$\frac{\pi((q) \to (q'))\, \rho((q) \to (q'))}{\pi((q') \to (q))\, \rho((q') \to (q))} \quad \exp(\quad \beta \Delta E) \tag{3}$$

with the change in energy induced by the proposed move, $\Delta E \quad E(q') \quad E(q)$. For uncorrelated random moves drawn from a given distribution (e.g., *x*, *y*, *z* coordinates or orientation of a particle, internal degrees of freedom) where the probability to construct the move is equal to the probability to construct the inverse move, $\pi((q) \to (q')) \quad \pi((q') \to (q))$, the probabilities of the move construction in eq. (3) cancel out. From this, Metropolis et al.[16] derived the Metropolis acceptance criterion:

$$\rho((q) \to (q')) \quad \begin{cases} \exp(\quad \beta \Delta E), & \Delta E > 0 \\ 1, & \Delta E \leq 0 \end{cases} \tag{4}$$

which is widely used for the simulation of physical and chemical systems.

While the construction of such moves is conceptually and technically simple, it often leads to inefficient simulation protocols. Consider a system consisting of *N* particles with a total of 3 *N* degrees of freedom. If only one or O(1) degrees of freedom are changed in one move, the algorithm will become inefficient in the thermodynamic limit and underperforms in comparison to MD simulations which change all degrees of freedom simultaneously. The evolution of the system in the thermodynamically relevant ensemble often requires collective changes of degrees of freedom which are exponentially unlikely to be proposed as a sequence of "local" moves. This problem arises both when the moves are proposed in sequence (which is often better in practice) or as collective moves composed of uncorrelated "local" moves. For many relevant potentials, the acceptance rate will decrease exponentially with the number of uncorrelated "local" moves in the collective move. In either scenario, the acceptance probability is low and the autocorrelation time high.

If the existence of steep energy rises in local moves is limiting the autocorrelation time of the method, one should aim to construct collective moves that reduce the fraction of such events in the move construction. We note that the Metropolis algorithm is only one particular choice to realize the more general condition of eq. (3). An alternative interpretation of eq. (3) is to postulate an ideal algorithm where each proposed change in the configuration will be accepted, that is, $\rho((q) \to (q')) \quad \rho((q') \to (q)) \quad 1$. In such an algorithm, detailed balance must be obeyed by constructing the probabilities π to select a move from all possible moves. In such an algorithm, a proposed moved has to be constructed such that

$$\frac{\pi((q) \to (q'))}{\pi((q') \to (q))} \quad \exp(\quad \beta \Delta E). \tag{5}$$

To exactly satisfy this criterion, the energy difference ΔE of all considered moves has to be known for all possible moves before each MC step to compute the probabilities $\pi((q) \to (q'))$. In the Swendsen–Wang family of methods, this is possible due to the local nature of the Hamiltonian, but this type of approach is unrealistic for most systems with continuous degrees of freedom and long-range potentials. For processes with a time-independent energy function, we can assume $\Delta E((q') \to (q))$ $\Delta E((q) \to (q'))$, (which is not true for example for metadynamics simulations[38–40] where an additional bias is applied to the system depending on its prior trajectory and the energy model thus changes with simulation time) and the relationship in eq. (5) is met by proposing moves with a probability:

$$\pi((q) \to (q')) \quad \frac{1}{A} \exp\left(\quad \beta \frac{\Delta E}{2} \right) \tag{6}$$

where $\Delta E \quad E(q') \quad E(q)$ is the change in energy by moving from q to q' and $A \quad \sum_{q'} \exp\left(\quad \beta \frac{\Delta E}{2} \right)$ is the normalization factor.

To illustrate this idea, let us consider an idealized system where the energy change ΔE is known for every possible move. The system consists of many particles where every particle has only two possible states, for example, the adsorption of a gas of independent particles on a surface. In this case, there are only two possible states and moves (from the adsorbed state into the gaseous state and vice versa) for which the energy change is known exactly a prior the move. Let us assume that the adsorbed state is energetically favorable by an energy $\Delta E_0 > 0$. Then according to eq. (6), the probability to adsorb a particle from the vacuum onto the surface is

$$\pi_a \quad \frac{1}{A} N_v \exp\left(\frac{1}{2} \beta \Delta E_0 \right), \tag{7}$$

while the probability to release a particle is

$$\pi_v \quad \frac{1}{A} N_a \exp\left(\quad \frac{1}{2} \beta \Delta E_0 \right), \tag{8}$$

where N_a and N_v are the number of particles in the adsorbed and vacuum state, respectively, and $A \quad N_v \exp\left(\frac{1}{2} \beta \Delta E_0\right) + N_a \exp\left(\quad \frac{1}{2} \beta \Delta E_0 \right)$ is the normalization factor.

In the standard Metropolis MC protocol, a particle is selected at random, and a move to the other state is proposed and accepted or rejected according to eq. (4). This leads to the rejection of most moves of particles which are already adsorbed on the wall, as $\Delta E > 0$. The biased protocol conversely picks a particle according to the Boltzmann probability of the energy change induce by its change of state and every move is accepted.

For realistic systems, the computational costs to compute the exact change in energy for every possible move are unacceptably high. However, a suitable estimator for the energy difference $\widetilde{\Delta E} \approx \Delta E$ can be used to construct the move probabilities and accelerate the performance of the method. For systems with classical interactions (e.g., molecular or atomistic systems), the obvious approximation is the Taylor expansion of the energy term. Using the force **F** and displacement $\Delta \mathbf{x}$, the change in energy can be approximated by

$$\widetilde{\Delta E_i} \approx \nabla E|_x \, \Delta \mathbf{x} \qquad \mathbf{F}(x)\Delta \mathbf{x} \tag{9}$$

and for the backward move, we have $\Delta \mathbf{x}'$ $\Delta \mathbf{x}$ and

$$\widetilde{\Delta E'} \approx \nabla E|_{x'} \, \Delta \mathbf{x}' \qquad \nabla E|_{x'} \, \Delta \mathbf{x} \quad \mathbf{F}(x') \, \Delta \mathbf{x} \tag{10}$$

This gives probabilities for the construction of move and back move: Inserting this into eq. (6) and then eq. (3) delivers the modified acceptance criterion:

$$\frac{\rho}{\rho'} \quad \frac{A}{A'} \exp\left(\frac{1}{2}(\nabla E|_x + \nabla E|_{x'}) \, \Delta \mathbf{x}\right) \exp(\quad \Delta E) \tag{11}$$

Here, A and A' are the sums over all probabilities for the moves and back moves. If more than one degree of freedom is modified during one MC step, the acceptance criterion becomes:

$$\frac{\rho}{\rho'} \quad \left(\frac{A}{A'}\right)^m \left[\prod_{i=1}^{m} \exp\left(\frac{1}{2}\left(\nabla E|_{x_i} + \nabla E|_{x_i'}\right) \Delta \mathbf{x}_i\right)\right] \exp(\quad \Delta E) \tag{12}$$

where m is the number of simultaneous MC moves.

This acceptance criterion asserts detailed balance exactly, as the energy change estimated by the Taylor expansion during the move construction [eqs. (6) and (9)] is corrected by the real change in energy.

AROMoCa approach. As explained in the previous section, the simultaneous random change of all degrees of freedom in large systems inevitably leads to vanishingly low acceptance rates in standard MC simulations. In the following, we propose an algorithm that first identifies the regions of the system that are far from the equilibrium and proposes moves to such regions with higher probability. We focus the derivation on the description of molecular systems, as most MD/MC simulations are performed for such systems. Without loss of generality, we consider a system with only translational degrees of freedom (e.g., a system consisting of interacting spherical particles) for simplicity. It is simple to extend this algorithm to other degrees of freedom, such as rigid rotations or changes applied to internal degrees of freedom of molecules.

Each MC steps is divided into four parts: First, for each degree of freedom (in our case: every particle and every direction), a probability to perform a displacement of any size between $\Delta x_{\max}$, $\Delta x_{\max}$ is calculated:

$$\pi_{i,j} \quad \frac{1}{A} \int_{\Delta x_{\max}}^{\Delta x_{\max}} dx \exp\left(\frac{F_{i,j}\, x}{2\, k_B T}\right) \tag{13}$$

$$\frac{1}{A} \frac{2\, k_B T}{F_{i,j}} \left[\exp\left(\frac{F_{i,j}\, \Delta x_{\max}}{2\, k_B T}\right) \quad \exp\left(\quad \frac{F_{i,j}\, \Delta x_{\max}}{2\, k_B T}\right)\right] \tag{14}$$

with a normalizing factor

$$A \quad \sum_i \sum_j \int_{\Delta x_{\max}}^{\Delta x_{\max}} dx \exp\left(\frac{F_{i,j}\, x}{2\, k_B T}\right) \tag{15}$$

Here, i denotes the particle index and j $0, 1, 2$ the direction index for movement in x, y, and z direction. The larger the force acting on a particle in a certain direction the larger $\pi_{i,j}$ for this specific move.

In a second step, $m\,(m \in N)$ moves are picked from the probability array constructed of all $\pi_{i,j}$. Neither are moves picked completely randomly nor are all degrees of freedom changed simultaneously. On the contrary: degrees of freedom that are far from equilibrium having a large force and large value of $\pi_{i,j}$ are more likely to be changed than those near local equilibrium. After identifying the m different particle-direction combinations, the displacement $\Delta x_{i,j}$ $\Delta x_{\max}\, \xi_{i,j}$ is calculated by drawing $\xi_{i,j}$ from the probability distribution:

$$\pi\left(\xi_{i,j}\right) \quad \frac{1}{B} \exp\left(\frac{F_{i,j}\, \xi_{i,j} \Delta x_{\max}}{2\, k_B T}\right) \quad \frac{1}{B} \exp\left(\frac{F_{i,j}\, \Delta x_{i,j}}{2\, k_B T}\right) \tag{16}$$

which is the equivalent of eq. (6). Again B $\int_{\Delta x_{\max}}^{\Delta x_{\max}} dx \exp\left(\frac{F_{i,j}\, x}{2\, k_B T}\right)$ is the normalization factor for this probability. It also appears in eqs. (13) and (14) and cancels out when calculating the total move probability as the product of $\pi_{i,j}$ and $\pi\left(\xi_{i,j}\right)$.

Numerically, the calculation of $\xi_{i,j}$ is performed by drawing a number η between 0 and 1 from a uniform random distribution and calculating $\xi_{i,j}$ using

$$\Delta x_{i,j} \quad \frac{2\, k_B T}{F_{i,j}} \ln\left(\eta \left(\exp\left(|\gamma_{i,j}|\right) \quad \exp\left(\quad |\gamma_{i,j}|\right)\right) + \exp\left(\quad |\gamma_{i,j}|\right)\right) \tag{17}$$

with $\gamma_{i,j}$ $\frac{F_{i,j}\, \Delta x_{\max}}{2\, k_B T}$. It can be shown that this results in the correct probability distribution in eq. (16) for the displacement.[29] After displacing each (i,j)-combination, the forces $F_{i,j}{}'$ and the normalization factor A' of the new configuration are calculated. The new configuration is then accepted according to eq. (12). All simulations reported below were performed with a Python-based implementation of AROMoCa using the OpenMM package[41] for the calculation of forces and energies between the atoms.

Results

In the following, we will present the sequence of investigations of increasingly complex systems that show the advantages and limitations of the AROMoCa approach. As the computational cost of a single step of the simulation of molecular systems has different contributions that are dependent on system size and difficult to compare the efficiency of the algorithms will be measured in terms of number simulation steps, which does not depend on the hardware used. In molecular simulations, the major contribution to simulation time is the evaluation of the long-range particle–particle interactions. As those interactions (forces and energies) have to be calculated for every step in all methods, the additional cost imposed by the AROMoCa algorithm is only O(N), corresponding to the evaluation of the estimated acceptance probabilities for the possible moves. This is typically much smaller (<10%) of the cost of the energy/gradient evaluation, so that the number of steps comparism gives a good estimate of the performance of the AROMoCa approach.

Independent particles in a double-well potential

In the first set of simulations, we test the efficiency of AROMOCA with respect to the movement of many particles at the same time. As discussed above, the efficiency of MC as compared with MD decreases rapidly with system size even for systems where the energy landscape is simple if only single particles are moved. We, therefore, investigated a system of noninteracting particles (here $N = 10^4$ initially placed with the positions uniformly distributed in an interval between $x_{min} = -10.0$ nm and $x_{max} = 10.0$ nm) in a double-well potential given by

$$U_{DW} = c_0 \left(a x^4 + b x^2 + c x\right) \quad (18)$$

with $a = 0.0005/\text{nm}^4$, $b = -0.04/\text{nm}^2$, and $c = 0.05/\text{nm}$ (see inset Fig. 1a). The potential strength was set to $c_0 = 5.0\, k_B T$.

The system was chosen because its dynamics comprises two phases: in the first phase, each particle will quickly relax to its nearest local minimum, but the distribution of the particles amongst the minima (initially approx. 50:50) will be far from the thermodynamic equilibrium of the system. The system will, therefore, equilibrate on the second, much slower time-scale, which is determined by the barrier between the two minima, and this relaxation is characterized by the rate by which particles cross the barrier from either the left or the right. In any method that reaches thermodynamic equilibrium, these rates will equal after equilibration, but different methods with differ in their absolute rates. The key goal of the development of accelerated simulations protocols is, therefore, the increase of the total rate, measured as function of the number of energy evaluations, which will control the accuracy of the determination of thermodynamic expectation values.

We performed AROMoCa and GMC simulations with initially equidistributed particles comprising 10×10^6 MC steps with a fixed displacement of 1.0 nm to either the right or left and $k_B T = 1.0$ kJ/mol. The height of the barrier in this system is 2.54 $k_B T$ for particles in the left (lower) minimum and 5.70 $k_B T$ for particles in the deeper right minimum, and 13 consequtive steps are required to move a particle from one minimum to the other. We performed AROMOCA simulations with collective moves for $m = 1, 2, 4, 8, 16$, and 32 independent particle displacements per step. The average energy per particle as a function of the number of steps is illustrated in Figure 1a. The two phases of the simulation are clearly visible in the slope of the energy relaxation. GMC and AROMoCa with $m = 1$ show approximately equal energy relaxation whereas for $m \geq 2$ AROMoCa converges much faster in energy (up to a factor of 20 for $m = 32$) than GMC.

We have computed the final occupation of distribution of particles and compared it to the Boltzmann distribution in the given potential. As expected both methods converge to the Boltzmann distribution in the long-time limit (inset of Fig. 1b),

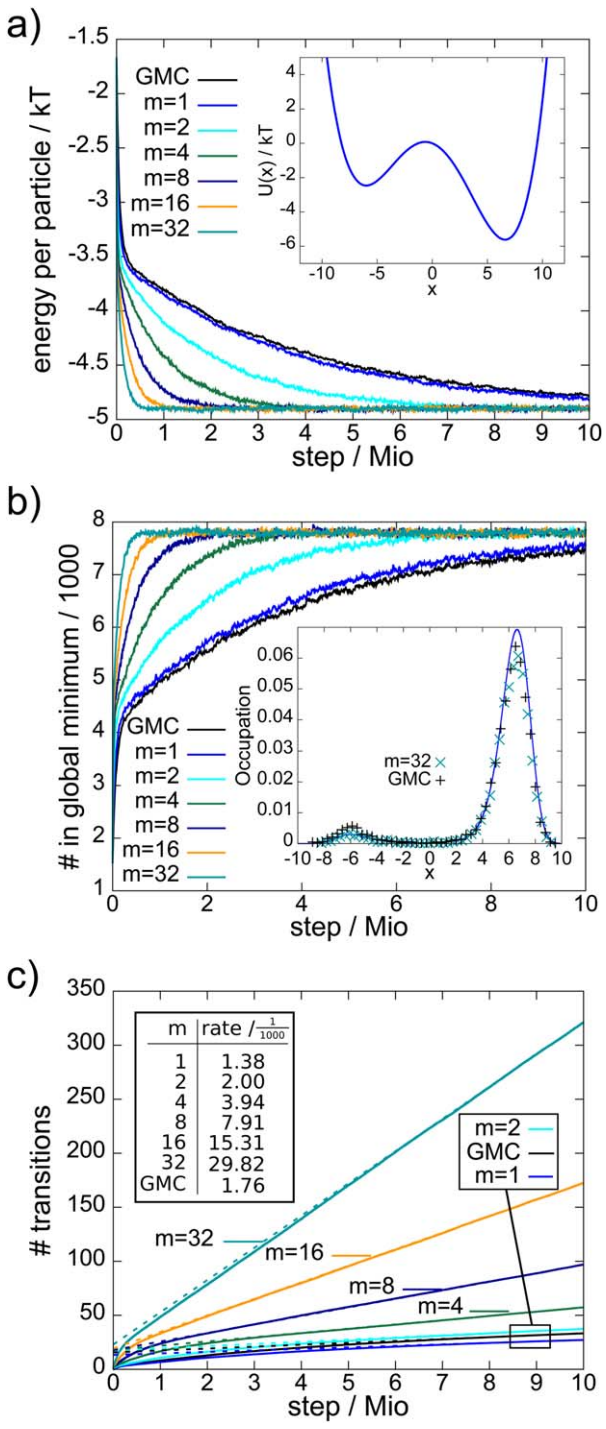

Figure 1. Comparison of generic Metropolis MC (GMC) and AROMoCa simulations with $m = 1, 2, 4, 8, 16$, and 32 multiple displacements per MC step. a) Energy convergence in an one dimensional double well potential (inset). While AROMoCa with $m = 1$ and GMC perform equally, equilibration accelerates with higher m and the equilibrium state is reached up to 20 times faster using AROMoCa with $m = 32$ than in the GMC simulation. b) The number of particles that reach energies within $k_B T$ of the global optimum converges faster in AROMoCa for larger values for m. The inset shows the final distribution function of the particles after the 10^7 MC steps in GMC and AROMoCa ($m = 32$) simulations compared with the exact distribution function. c) As the particles were initially equidistributed between $x_{min} = -10.0$ and $x_{max} = 10.0$, the convergence is measured by the transition rate of particles across the barrier. The particle current across the energy barrier at $x = 0$ is shown as a function of the number of simulation steps. A linear function is fitted to the transition values of the last 2×10^6 MC steps to determine the transition rates (table inset). The results clearly show a strongly increased relaxation rate in the system for AROMoCa with $m \geq 2$ in comparison to GMC, leading to an overall increased performance. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

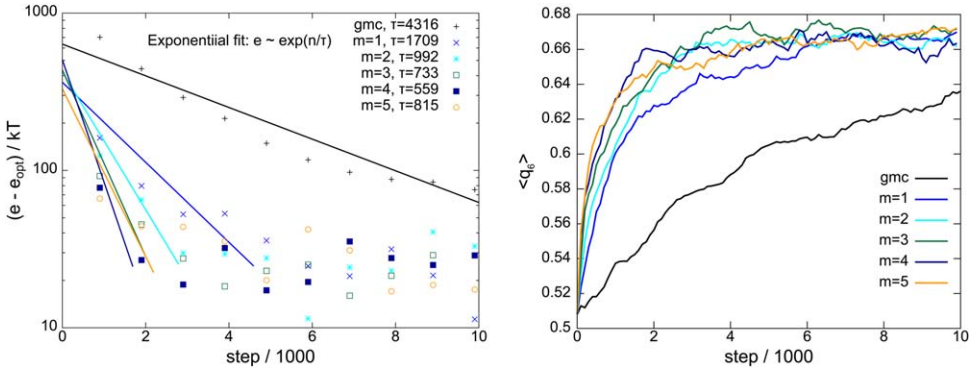

Figure 2. Liquid to crystal transition of a two dimensional Lennard Jones liquid. The convergence speed in terms of total system energy (left) and the q_6 bond order parameter (see text) averaged over all particles (right) of generic Metropolis MC is compared with AROMoCa using different m. Exponential fits of the decay in energy show that AROMoCa converges up to six times faster than generic MC. The rate of convergence increases with the number of simultaneous displacements per MC step m for up to m 4. The trends of the q_6 bond order parameter verify that the total order of the system is increased throughout the simulations faster in the AROMoCa simulation than using GMC. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

but the AROMOMA simulations converge much faster, even though the particles are independent, that is, there are no collective effects. Figure 1b shows the fraction of particles that are within k_BT of the global optimum. As for the energy per particle, plotted in Figure 1a, AROMoCa with $m \geq 2$ performs much better than GMC. As explained above, the sampling speed of the algorithm is determined by the total transition rate of particles across the barrier. Figure 1c shows the sum of the left-to-right and the right-to-left particle currents crossing the barrier at approximately x 0. The currents determined from linear fits to the last 2×10^6 MC steps are given in the inset. The transition rate of AROMoCa increases with nearly unit slope almost linearly with m, which results in a much more efficient algorithm for $m \geq 2$.

Recalling the AROMOCA algorithm, this appears counterintuitive at first: Particles that are not in equilibrium are much more likely to be selected and moved by AROMOCA than particles near the minimum and the move selection criterium [left or right, see eqs. (13) and (14)] will favor relaxation to the nearest local minimum and not crossing of the barrier. AROMOCA nevertheless outperforms GMC because of its high acceptance rate (inset table Figure 1c). As AROMOCA moves particles according to detailed balance, the local relaxation (moving with the gradient of the potential) must take precedence over barrier crossing. However, as many particles are moved with near unit acceptance rate in every single step, the relaxation dynamics of the system is faster. This is demonstrated by comparing AROMOCA with m 32 with an MC protocol that moves 32 particles are random with the standard acceptance rate: The AROMOCA current across the barrier approximately 30 times higher in comparison.

Liquid-crystal transition of Lennard-Jones systems. To demonstrate the behavior of the algorithm for a system of interacting particles with hard-core potentials, AROMoCa was applied to a two-dimensional Lennard-Jones model for Argon (Lennard-Jones parameters ε 0.99601 kJ/mol and σ 0.3405 nm) at T 50 K which is below its melting point. Initially, 526 particles were distributed randomly on a square surface of 9.5×9.5 nm^2. The system size was chosen based on initial estimates on the convergence time for GMC. This system was chosen because the domain growth can be easily visualized; a three-dimensional (3D) model is studied in the next section. We performed GMC and AROMoCa simulations with m 1, 2, 3, 4, and 5 to compare the methods. In all simulations, the maximal displacement was set to $\Delta x_{\max}$ 0.75 nm. To evaluate the convergence of AROMoCa and GMC, we monitored the total system energy per step and the q_6-bond-order parameter based on the parameter proposed by Steinhardt et al.,[42] which is defined as:

$$q_l(a) \quad \sqrt{\frac{4\pi}{2l+1}\sum_{m=\ l}^{l}\left|\frac{1}{n(a)}\sum_{NN(a)} f_c(r)\, Y_{lm}(\theta,\phi)\right|^2} \qquad (19)$$

where $n(a)$ 6 is the number of nearest neighbors $(NN(a))$ of atom a, (r,θ,ϕ) are the spherical coordinates of the vectors from atom a to the nearest neighbors, and Y_{lm} the spherical harmonics. The cutoff function

$$f_c \quad \left(1+\exp\left(\frac{r \quad r_0}{t}\right)\right)^{\ 1}$$

with t 0.15Å and r_0 5.451Å (1Å larger than the equilibrium distance) was used to ensure that the six nearest neighbors contribute to the order parameter only when within a certain limit.

Figure 2 shows the dependence of the average system energy and bond-order parameter. The energy relaxation in the liquid-to-crystal transition is well represented by an exponential function, $E(\text{step}) \propto \exp(\ \text{step}/\tau)$. The results of the fit, also displayed in Figure 2, show that the decay in energy occurs up to eight times faster in AROMoCa than in GMC (τ 559 vs. τ 4316). The increase in the local order of the system induced by crystal transition is demonstrated by the increase in the bond-order parameter in the right hand side of Figure 2. As for the energy, the order parameter increases much faster using AROMoCa than GMC. To illustrate the ordering of the system, Figure 3 shows snapshots at step 0, 100, 1000, 2500, and 5000 of the GMC trajectory and the trajectory of AROMoCa with

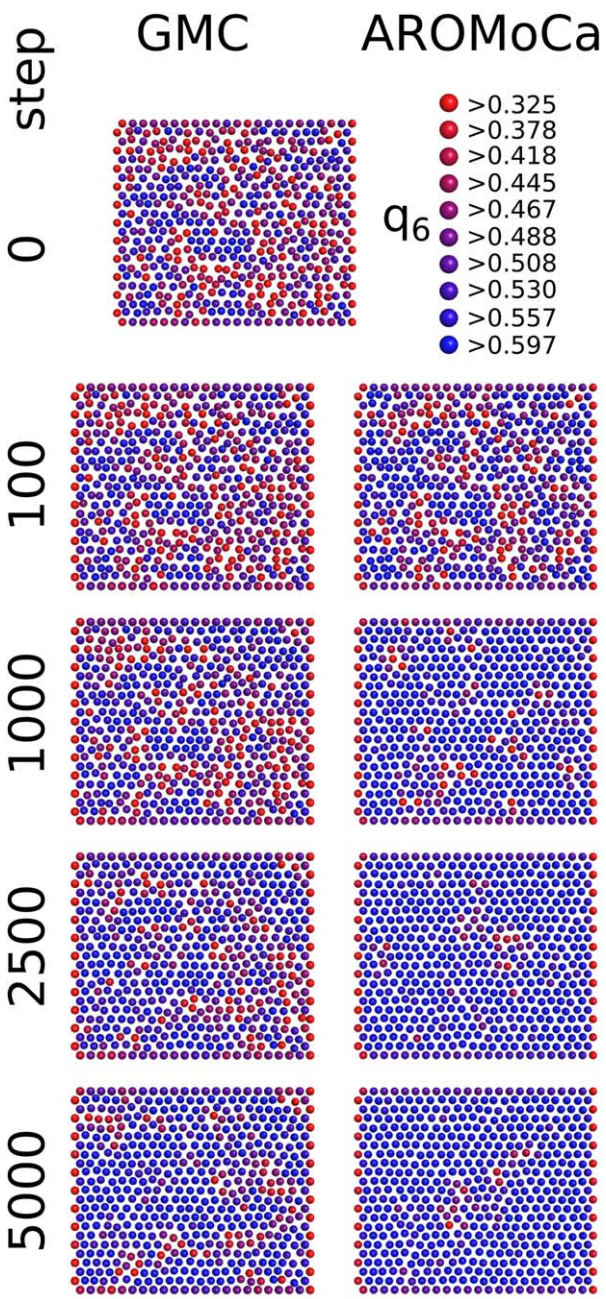

Figure 3. The q_6 bond order parameter of each atom and step was calcu lated to analyze the behavior of the 2D Lennard Jones model throughout the simulation. The color code indicates the degree of order for each parti cle as indicated in the top right. Snapshots for steps 0, 100, 1000, 2500, and 5000 are displayed. First, in AROMoCa, the system reaches a nearly crystal line configuration after less than 2500 steps and defects (red particles) occur afterwards to a constant fraction due to thermal fluctuations, while GMC does not reach the same degree of crystallinity after 5000 MC steps.

m 4, respectively. The Figure 3 clearly demonstrates that the AROMoCa simulation reduces the fraction of disordered atoms, that is, the particles not in crystalline environment, rapidly to a constant fraction that arises from thermal fluctuations within the first 2500 MC steps, while GMC has not yet reached the same degree of order after 5000 MC steps.

To further analyze the mechanism of the ordering process, we estimated the growth of crystalline domains in the sample. For every snapshot (snapshots recorded every 100th simulation step over a simulation of 10k MC steps), the distribution of the distance ζ of each atom to the nearest "disordered" particle (defined by $q_6 < 0.418$), based on the values of q_6 occurring in the simulation, that is, the nearest particle not in a crystalline environment was measured. Figure 4 illustrates the definition of ζ, its distribution, and the evolution of its mean value $<\zeta>$. As fluctuations were quite large due to the small size of the system, we computed a moving average over up to 20 subsequent frames of the simulation (a lower number was averaged in the first 20 frames). This trend again illustrates the advantage of AROMoCa over GMC: $<\zeta>$ reaches an equilibrium value of 40 Å, which is just below half the system size, much faster for all values of m in AROMoCa in comparison to GMC. Furthermore, the performance of AROMoCa improves with higher m.

Finally, we studied a 3D LJ liquid consisting of 216 particles and compared the results of AROMoCa and GMC simulations with a MD simulation of the same system. AROMoCa simulations were performed using m 1, 2, and 4 displacements per MC steps with a maximal step size of Δx_{max} 0.75 nm. We again used Argon parameters and a temperature of T 50 K. The MD simulation was performed using the Langevin integrator as implemented in the OpenMM toolkit[41] with a time step of Δt 1.0 fs. Periodic boundary conditions were applied in all three dimensions in all simulations. The system energy was recorded throughout the simulations and is plotted in Figure 5a as a function of the number of simulation steps. In the MD and AROMoCa simulations for m 4, we observe two plateaus in the energy, corresponding to the supercooled liquid and crystal, respectively. This is demonstrated by the radial distribution functions shown in Figure 5b for the two regions in the left panel of the Figure 5a. Analysis of the radial distribution functions indicates that all simulations converge to a supercooled liquid state (encircled by dashed lines, dashed RDF plots) before the crystal transition takes place. For those trajectories where the transition can be observed the RDF change their shape and show additional distinct peaks (encircled in solid lines, solid RDF plots). While no phase transition occurred in the GMC runs and with AROMoCa using $m \leq 2$, it was observed for the MD simulation and for the AROMoCa simulation with m 4 between the simulation steps 4×10^5 and 6×10^5 at approximately the same number of function evaluations. This indicates that AROMoCa can model some collective processes with comparable computational efficiency as MD simulations.

Detailed balance and acceptance rates: Limits of fbMC

In the following section, we aim to compare AROMoCa with a similar approach to introduce gradient-driven collective moves in MC methods, in particular with force-bias MC methods.[30] fbMC and other previously developed methods[24,25,29,30] use the gradient acting on single particles to determine the step size by which every single degree of freedom is to be changed in one simulation step. The displacements are constructed in such a way that the ratio of the probability to perform a move to the probability to perform the back move is approximately

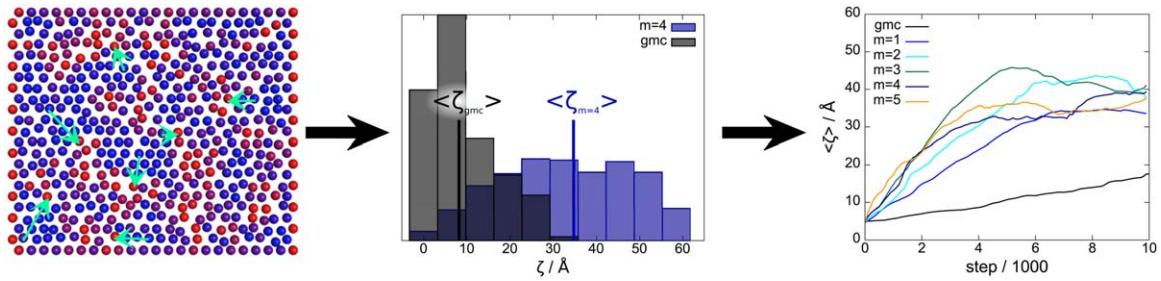

Figure 4. Analysis of the growth of crystalline areas in the 2D system: For each particle and step, ζ, the distance to the nearest particle that is not in a crystal line environment (defined by $q_6 < 0.418$) is measured (left). The central panel shows the distribution of the GMC simulation and the AROMoCa simulation using m 4 at step 5000 and the right panel the moving average of $<\zeta>$ for AROMoCa with m 1, . . . 5 and GMC. $<\zeta>$ reaches an equilibrium value of 40Å (just below half the system size) much faster with ARMoCa even when using m 1 than with GMC. Additionally AROMoCa performance improves with increasing m.

the Boltzmann weight, as $\Delta E \quad \nabla E|_x \, \Delta \mathbf{x} \quad \mathbf{F}(x)\Delta x$. Different versions of this approach have been proposed: In some methods, all proposed steps are accepted under the assumption that detailed balance is only weakly violated, while others retain an acceptance criterion. In the first approach, the acceptance rate is unity by definition, but thermodynamic expectation values may be incorrectly computed, while the efficiency of the second approach depends on the acceptance rate for a particular system, just as in any other MC method. In the following, we, therefore, investigated both variants of the fbMC method in comparison with AROMoCa for simple test system to determine the impact of an acceptance criterion on thermodynamic expectation values, such as the mean energy of the ensemble and on the acceptance rates.

We consider a small system of Argon with 216 atoms with Lennard-Jones interactions. First, the system was equilibrated at three different temperatures (T 200 K, T 80 K, and T 40 K) to construct ensembles in gaseous, liquid, and solid phase. After the equilibration, AROMoCa and fbMC simulations with 10^6 MC steps each were performed at each temperature using random displacements in all three dimensions with step sizes ranging from 0.005 nm to 0.075 nm. As a reference, a 100 ps MD run at each temperature was performed. The averaged energies in the MD runs are 1715.1±7.0 kJ/mol, 1192±11 kJ/mol, and 165.7±9.8 kJ/mol for solid, liquid, and gaseous system respectively. The mean energy and the standard deviation plotted over the step size are displayed in Figure 6.

In the gaseous and liquid systems, the energies of the ensemble simulated with the fbMC method resemble the reference energy for small steps but show an increasing deviation with steps larger than 0.025 nm: In the gaseous system, the relative deviations from the energy averaged over the MD simulations are between 1.2 and 3.3% for the step sizes between 0.01 and 0.025 nm but 25% and even 155% for step sizes of 0.05 and 0.075, respectively. In the liquid system, the relative deviation is lower than 2.5% for step sizes of up to 0.025 nm but already 50% for step sizes of 0.05 nm. In the simulation of the condensed system, fbMC without acceptance criterion completely fails to represent the physical ensemble and the expectation values of the energy increase drastically with the step size: although the relative deviation is less than 4% for step sizes of up to 0.025 nm, the absolute difference of about 7.166 kJ/mol even for the smallest step of 0.01 nm is more than 21 times $k_B T$ at the simulation temperature of 40 K. This indicates that the approximation of the change in energy by the product of gradient and displacement yields adequate results only for very small step sizes and points out the necessity to use an acceptance criterion in fbMC to obtain proper thermodynamic averages. AROMoCa with the acceptance criterion derived in eq. (12) conversely manages perfectly to simulate physical ensembles with the correct energies for all temperatures and step sizes up to 0.075 nm: the highest deviation of all step sizes in all systems is lower than 1% and the absolute deviation in the crystalline system is only 1.55 kJ/mol for the largest tested step size of 0.075 nm.

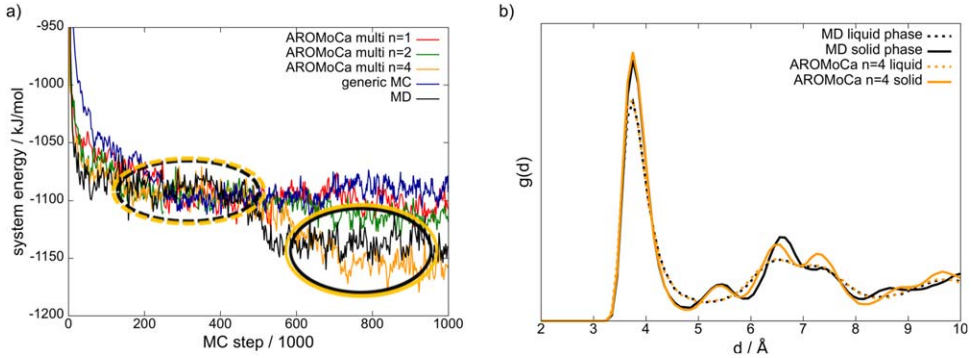

Figure 5. Liquid to crystal transition of a 3D Lennard Jones liquid. The radial distribution functions show distinct peaks after the transition process (solid lines) in comparison to the supercooled liquid phase (dashed lines) indicating a crystalline phase. The transition to a crystal is not observed at all in the standard Metropolis MC simulation. AROMoCa relaxes the system with approximately the same rate as the MD simulation. The right panel shows radial distribution functions for the encircled regions in Figure 5a by averaging over 200 frames of the trajectories. The radial distribution functions were plotted for the MD run and the AROMoCa run with m 4. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

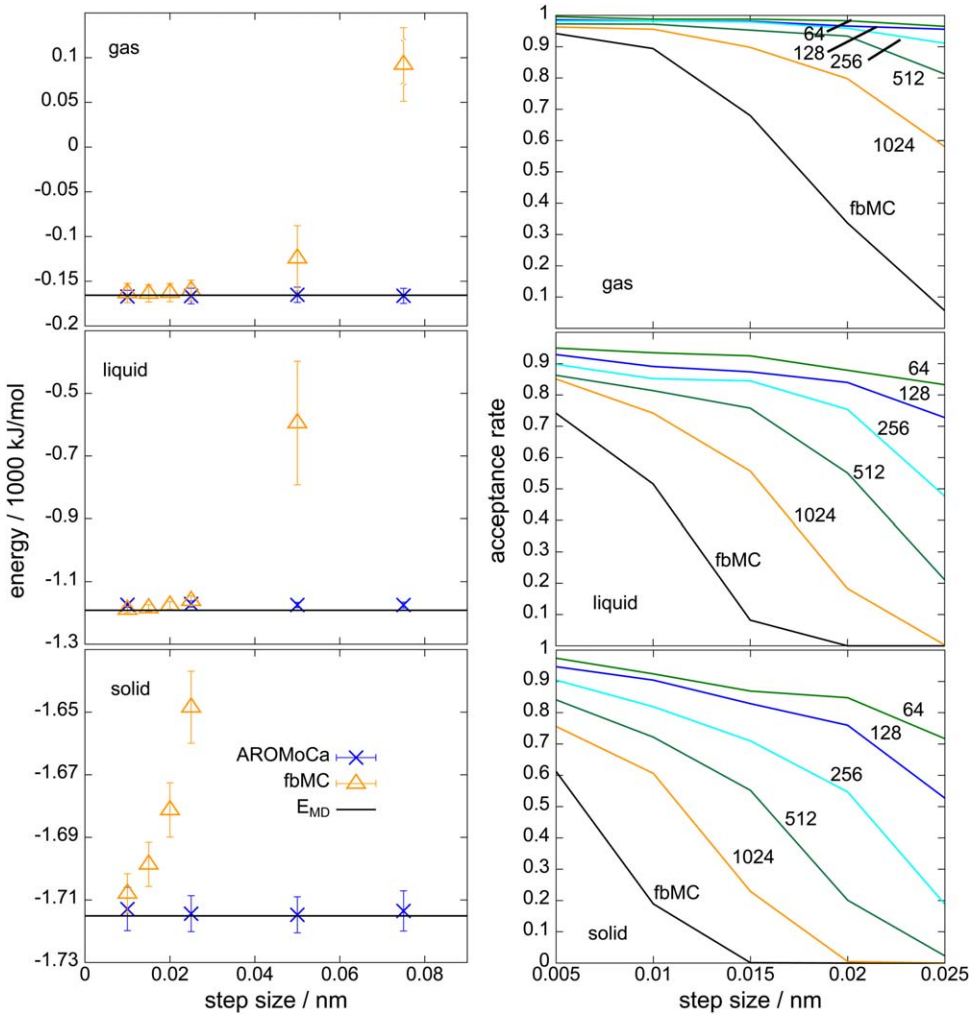

Figure 6. Analysis of detailed balance and acceptance rates of fbMC and AROMoCa in different phases of the simulations. Left: Energy averaged over 10^5 snapshots of a 10^6 step simulations of an Argon system in the gaseous (T 200 K, top), liquid (T 80 K, middle), and solid state (T 40 K, bottom) for different methods (AROMoCa and fbMC) as a function of the step size. No acceptance criterion was applied to the fbMC simulations. MD simulations of the same systems were performed for comparison. AROMoCa reproduces the MD energy distributions (-1715.1 ± 7.0 kJ/mol, -1192 ± 11 kJ/mol, and -165.7 ± 9.8 kJ/mol, respectively) very well. Right: Acceptance rates of AROMoCa and fbMC runs with 1000 MC steps of gaseous, liquid, and solid Argon (1728 particles at T 200 K, 80 K, and 40 K from top to bottom) for different step sizes. AROMoCa simulations were performed using m 64 1024 system changes per MC step. Especially, for larger step sizes (>0.1 nm), AROMoCa reaches much higher acceptance rates than fbMC. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

To test the influence of an acceptance criterion applied in fbMC, a system consisting of 1728 Lennard-Jones particles (same parameters as above) was studied. For the same temperatures as above (T 200 K, T 80 K, and T 40 K) simulations consisting of 1000 MC steps were performed after initial equilibration at each temperature. For the fbMC simulation, the acceptance criterion of eq. (12) was applied with the product over all particles and directions. The acceptance rates were calculated for step sizes between 0.005 and 0.025 nm. AROMoCa simulations were performed for m 64, 128, 256, 512, and 1024 parallel displacements per MC step. The resulting acceptance rates are displayed in the right hand side of Figure 6. For the gaseous state, the AROMoCa acceptance rates are close to 1.0 for up to m 512 parallel displacements for all step sizes. The fbMC acceptance rate becomes vanishingly small for step sizes close to 0.025 nm. In the liquid and solid phase acceptance rates larger than 0.5 can be achieved in AROMoCa by limiting the movement to the most important regions of the system. The smaller m, the larger the probability to only move those parts of the system that are far from the equilibrium. This increases the probability to improve the total system energy and leads to high acceptance rates. Because fbMC cannot differentiate between the regions of the system close to equilibrium and the regions farther away from equilibrium the acceptance rate is reduced drastically.

Conclusions

Molecular simulations are presently limited in system size and time-scale by the computational effort required to overcome long autocorrelation times. In this work, we investigated the AROMoCa algorithm as a generic approach to generate complex collective moves with high acceptance rates in MC simulations. By performing changes to the system based on an

estimator for the change in energy induced by a potential move, higher acceptance rates can be achieved. It was shown that AROMoCa converges important physical observables, such as order parameters, faster than Metropolis MC methods for all of the systems studied in this work. For the simulations of the crystallization of a 3D Lennard-Jones liquid, the convergence of AROMoCa was comparable to MD simulations of the same system. A similar protocol, which moves a single particle, has been previously investigated for the equilibration of water.[24] Here, we find that moving a single particle is often insufficient to significantly accelerate the simulation convergence, but that a balance between the number of particles moves and the step size must be found to obtain an optimal algorithm.

The computation of gradients has efficiently been implemented in many MD programs and incurs roughly a factor of three in computational effort compared with the energy evaluation alone. The total computational effort in AROMoCa is, therefore, still dominated by the evaluation of gradients and energy, the cost of move construction was below 10% in the simulations reported here.

In summary, we found that AROMoCa is a generic MC algorithm that can accelerate the simulation of many systems. It can be combined with or embedded in other methods such as MTM,[21,22] SA,[20] parallel tempering,[19] or Model Hopping methods[43] that presently are based on generic Metropolis MC methods. This opens a variety of possibilities for improvement of computational efficiency for many applications. Although similar ideas have already been explored in fbMC (force biased Monte Carlo) or uniform-acceptance fbMC,[29,30] previously established methods lack one key feature: the identification of the degrees of freedom with values far from the local equilibrium, which is the central element of the AROMoCa method, is the essential new ingredient that generates acceptance rates close to unity for collective moves, while exactly preserving detailed balance at all times.

Acknowledgments

The authors acknowledge support from the BMBF program ^pageRWid=177.8mmMolecular Interaction Engineering. Furthermore TN thanks the Ministry of Science, Research and Arts and the Universities of the State of Baden-Württemberg for funding by the Landesgraduiertenförderung. The authors gratefully acknowledge fruitful discussions with Mitsunori Nakamoto, Yuichi Tokita, Vadim Rodin, and Florian von Wrochem during the development of this method.

Keywords: Monte Carlo · molecular modeling · molecular simulation · unit acceptance probability

[1] A. R. Leach, *Molecular Modelling: Principles and Applications*, 2nd ed.; Pearson Education Limited: Harlow, **2001**.
[2] S. Sastry, P. G. Debenedetti, F. H. Stillinger, *Nature* **1998**, *393*, 554.
[3] Y. Sugita, Y. Okamoto, *Chem. Phys. Lett.* **1999**, *314*, 141.
[4] Y. Sugita, A. Kitao, Y. Okamoto, *J. Chem. Phys.* **2000**, *113*, 6042.
[5] A. Schug, T. Herges, W. Wenzel, *Phys. Rev. Lett.* **2003**, *91*, 158102.
[6] J. J. Kwiatkowski, J. Nelson, H. Li, J. L. Bredas, W. Wenzel, C. Lennartz, *Phys. Chem. Chem. Phys.* **2008**, *10*, 1852.
[7] A. Lukyanov, C. Lennartz, D. Andrienko, *Phys. Status Solidi (a)* **2009**, *2*, 2737.
[8] K. Lindorff Larsen, S. Piana, R. O. Dror, D. E. Shaw, *Science* **2011**, *334*, 517.
[9] K. A. Dill, J. L. MacCallum, *Science* **2012**, *338*, 1042.
[10] T. Strunk, M. Wolf, M. Brieg, K. Klenin, A. Biewer, F. Tristram, M. Ernst, P. J. Kleine, N. Heilmann, I. Kondov, W. Wenzel, *J. Comput. Chem.* **2012**, *33*, 2602.
[11] T. Neumann, D. Danilov, C. Lennartz, W. Wenze, *J. Comput. Chem.* **2013**, *34*, 2716.
[12] D. E. Shaw, M. M. Deneroff, R. O. Dror, J. S. Kuskin, R. H. Larson, J. K. Salmon, C. Young, B. Batson, K. J. Bowers, J. C. Chao, M. P. Eastwood, J. Gagliardo, J. P. Grossman, C. R. Ho, D. J. Ierardi, I. Kolossvýry, J. L. Klepeis, T. Layman, C. McLeavey, M. A. Moraes, R. Mueller, E. C. Priest, Y. Shan, J. Spengler, M. Theobald, B. Towles, S. C. Wang, In Proceedings of the 34th Annual International Symposium on Computer Architec ture; ACM, San Diego, California, **2007**; pp. 1.
[13] D. E. Shaw, R.O. Dror, J.K. Salmon, J.P. Grossman, K.M. Mackenzie, J.A. Bank, C. Young, M.M. Deneroff, B. Batson, K.J. Bowers, E. Chow, M.P. Eastwood, D.J. Ierardi, J.L. Klepeis, J.S. Kuskin, R.H. Larson, K. Lindorff Larsen, P. Maragakis, M.A. Moraes, S. Piana, Y. Shan, B. Towles, In Pro ceedings of the Conference on High Performance Computing Net working, Storage and Analysis; New York, NY, USA, **2009**; pp. 1.
[14] M. Tuckerman, B. J. Berne, G. J. Martyna, *J. Chem. Phys.* **1992**, *97*, 1990.
[15] J. A. Izaguirre, S. Reich, R. D. Skeel, *J. Chem. Phys.* **1999**, *110*, 9853.
[16] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, *J. Chem. Phys.* **1953**, *21*, 1087.
[17] J. S. Wang, R. H. Swendsen, *Phys. Rev. B* **1988**, *37*, 7745.
[18] J. S. Wang, R. H. Swendsen, *Phys. A Stat. Mech. Appl.* **1990**, *167*, 565.
[19] R. H. Swendsen, J. S. Wang, *Phys. Rev. Lett.* **1986**, *57*, 2607.
[20] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, *Science* **1983**, *220*, 671.
[21] J. S. Liu, F. Liang, W. H. Wong, *J. Am. Stat. Assoc.* **2000**, *95*, 121.
[22] R. Craiu, C. Lemieux, *Stat. Comput.* **2007**, *17*, 109.
[23] D. A. Kofke, *Mol. Phys.* **2004**, *102*, 495.
[24] C. Pangali, M. Rao, B. J. Berne, *Chem. Phys. Lett.* **1978**, *55*, 413.
[25] M. Rao, C. Pangali, B. J. Berne, *Mol. Phys.* **1979**, *37*, 1773.
[26] S. Whitelam, P. L. Geissler, *J. Chem. Phys.* **2007**, *127*, 154101.
[27] S. Whitelam, P. L. Geissler, *J. Chem. Phys.* **2008**, *128*, 219901.
[28] L. Velazquez, J. C. Castro Palacio, *Phys. Rev. E* **2013**, *88*, 013311.
[29] M. Timonova, J. Groenewegen, B. J. Thijsse, *Phys. Rev. B* **2010**, *81*, 144107.
[30] E. C. Neyts, B. J. Thijsse, M. J. Mees, K. M. Bal, G. Pourtois, *J. Chem. Theory Comput.* **2012**, *8*, 1865.
[31] G. Dereli, *Mol. Simul.* **1992**, *8*, 351.
[32] M. J. Mees, G. Pourtois, E. C. Neyts, B. J. Thijsse, A. Stesmans, *Phys. Rev. B* **2012**, *85*, 134301.
[33] K. M. Bal, E. C. Neyts, *J. Chem. Phys.* **2014**, *141*, 204104.
[34] J. Liu, E. Luijten, *Phys. Rev. Lett.* **2004**, *92*, 035504.
[35] W. Krauth, *Phys. Rev. Lett.* **1996**, *77*, 3695.
[36] M. P. Allen, D. J. Tildesley, *Computer Simulation of Liquids*; Oxford Uni versity Press: Oxford, UK, **1989**.
[37] K. A. Fichthorn, W. H. Weinberg, *J. Chem. Phys.* **1991**, *95*, 1019.
[38] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, M. Parrinello., *Comput. Phys. Commun.* **2009**, *180*, 1961.
[39] L. Alessandro, L. G. Francesco, *Rep. Prog. Phys.* **2008**, *71*, 126601.
[40] V. Limongelli, M. Bonomi, M. Parrinello, *Proc. Natl. Acad. Sci.* **2013**, *110*, 6358.
[41] P. Eastman, P. Eastman, M. S. Friedrichs, J. D. Chodera, R. J. Radmer, C. M. Bruns, J. P. Ku, K. A. Beauchamp, T. J. Lane, L. P. Wang, D. Shukla, T. Tye, M. Houston, T. Stich, C. Klein, M. R. Shirts, V. S. Pande, *J. Chem. Theory Comput.* **2013**, *9*, 461.
[42] P. J. Steinhardt, D. R. Nelson, M. Ronchetti, *Phys. Rev. B* **1983**, *28*, 784.
[43] W. Kwak, U. H. E. Hansmann, *Phys. Rev. Lett.* **2005**, *95*, 138102.