Optimizing replica exchange moves for molecular dynamics

Walter Nadler[1,*] and Ulrich H. E. Hansmann[1,2,†]
[1]*John-von-Neumann Institute for Computing, Forschungszentrum Jülich, D-52425 Jülich, Germany*
[2]*Department of Physics, Michigan Technological University, Houghton, Michigan 49931, USA*
(Received 27 August 2007; published 15 November 2007)

We sketch the statistical physics framework of the replica exchange technique when applied to molecular dynamics simulations. In particular, we draw attention to generalized move sets that allow a variety of optimizations as well as new applications of the method.

DOI: 10.1103/PhysRevE.76.057102 PACS number(s): 82.20.Wt, 87.15.Aa, 83.10.Rs, 02.70.Ns

Effective simulation of proteins [1], glasses [2], and similar complex systems [3] is hampered by slow relaxation due to barriers and bottlenecks. Replica exchange [4–6]—also known as parallel tempering—is one of the main approaches to overcome these problems. Originally devised for stochastic simulations, it is used nowadays also in combination with molecular dynamics (MD) simulations—i. e., simulations that have a strong deterministic character. We have found that there exists some confusion on the correct application of the replica exchange approach to MD. In this Brief Report we will sketch the theoretical basis for applying replica exchange to MD simulations, introduce generalized move sets necessary for optimizing exchange rates, and point to possible extensions of the concept.

In replica exchange, a set of stochastic simulations are performed in parallel with distinct weight functions $w_i(S)$; see Fig. 1. At certain times an exchange of current conformations S_i of replicas at neighboring samplers [7] is attempted, e.g., for a single pair of simulations

$$\{S_1,S_2\} \to \{S_1',S_2'\} = \{S_2,S_1\}. \tag{1}$$

Such an exchange move does not change the statistics of the full distribution function

$$w_{tot}(S_1,S_2) = w_1(S_1)w_2(S_2) \tag{2}$$

if it is accepted or rejected according to a generalized Metropolis rule [8]

$$p_M(\{S_1,S_2\} \to \{S_1',S_2'\}) = \min\left(1, \frac{w_{tot}(S_1',S_2')}{w_{tot}(S_1,S_2)}\right). \tag{3}$$

In physical and chemical applications one usually focuses on the canonical ensemble. Using replica exchange an individual replica performs a random walk in temperature, allowing it to enter and escape local potential minima. As a consequence, the state space is explored more thoroughly, especially at low temperatures. The weight functions employed in such situations are the canonical distributions $w(S) \propto \exp[-\beta E(S)]$, with $E(S)$ the energy of the system and $\beta = 1/k_B T$ the inverse temperature. For the exchange move of Eq. (1), the Metropolis criterion assumes the simple form

$$p_M(\{S_1,S_2\} \to \{S_2,S_1\}) = \min(1, \exp[\Delta\beta\Delta E]), \tag{4}$$

with $\Delta\beta = \beta_2 - \beta_1$ and $\Delta E = E(S_2) - E(S_1)$.

The exchange move of Eq. (1) is not unique. More general moves can be derived involving both exchange *and* some modification of the exchanged states:

$$\{S_1,S_2\} \to \{S_1'(S2), S_2'(S1)\}. \tag{5}$$

Such moves are allowed as long as they preserve detailed balance and—in combination with the independent simulations—do not violate global ergodicity. The exponent of the acceptance probability for such generalized replica exchange moves is no longer given by the simple form in Eq. (4) but by

$$\ln(W) \equiv \ln\left[\frac{w_{tot}(S_1',S_2')}{w_{tot}(S_1,S_2)}\right] = \beta_1[(E(S_1) - E(S_1')] + \beta_2[E(S_2) - E(S_2')\ ]. \tag{6}$$

Molecular dynamics—i.e., determining the time evolution of a classical many-body system by numerically solving the equations of motion—is an intrinsically deterministic approach. At first glance, this property seems to preclude its incorporation into the replica exchange scheme sketched above, since no stochastic sampling appears to be involved. However, already early on MD simulations were viewed also as a stochastic sampling procedure by considering the velocity field as a heat bath [10,11]: A continuous appropriate rescaling of the velocities leads to a correct canonical sam-

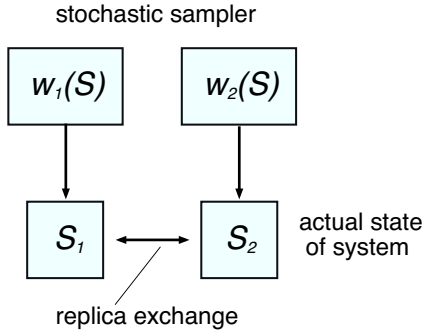


FIG. 1. Sketch of the basic situation in replica exchange simulations: simulations run independently on stochastic samplers with weight function $w_i(S)$; at certain times, exchange of the current conformations is attempted with probability p_M, Eq. (3).

*wnadler@mtu.edu, w.nadler@fz-juelich.de
†hansmann@mtu.edu, u.hansmann@fz-juelich.de

 ©2007 The American Physical Society

pling of all properties depending only on the coordinates of the particles [12,13]. In this way the system state is defined by the coordinates only, $S=\mathbf{x}$, and the velocity field in a way acts as stochastic sampler. The above replica exchange scheme can then be applied straightforwardly. Care has to be taken, however, to properly treat the velocity field as stochastic sampler also in this situation, e.g., like in Ref. [6] by randomly reinitializing the velocities upon exchange according to

$$\langle E_{kin}\rangle = \frac{1}{2}Nk_BT, \tag{7}$$

with N being the number of degrees of freedom.

By considering the velocity field as a heat bath one gives up the idea of a trajectory. For this and other [12] reasons, it is more common in canonical MD simulations to control the system by an additional thermostat algorithm [10,11,13]. In such an implementation of canonical MD the system state is given by coordinates *and* velocities together, $S=(\mathbf{x},\mathbf{v})$, the thermostat acting as the external stochastic sampler. Consequently, in replica exchange moves the full state of the system—i.e., $S=(\mathbf{x},\mathbf{v})$—has to be considered. The contributions to the energy can be separated as

$$E(S) = E_{pot}(\mathbf{x}) + E_{kin}(\mathbf{v}). \tag{8}$$

As a consequence, the acceptance probability of Eq. (6) now takes on the form

$$\ln(W) = \Delta\beta\Delta E_{pot} + \Delta K, \tag{9}$$

with

$$\Delta K = (\beta_2 - \beta_1)[E_{kin}(\mathbf{v}_2) - E_{kin}(\mathbf{v}_1)]. \tag{10}$$

Since the kinetic energy reflects the simulation temperature [see Eq. (7)], in general ΔK will be negative. Hence, a naive application of Eq. (1) is hampered by a large detrimental contribution from the kinetic energy difference yielding a very low acceptance probability, Eq. (4). Moreover, accepted moves also lead to velocities that are not characteristic for the new temperatures, pushing the system out of equilibrium.

One way to avoid such problems is to turn to generalized exchange moves that control the possibly large fluctuations in the velocity fields. A rescaling of the velocity fields in the move

$$\{(\mathbf{x}_1,\mathbf{v}_1),(\mathbf{x}_2,\mathbf{v}_2)\} \to \{(\mathbf{x}_2,r^{-1}\mathbf{v}_2),(\mathbf{x}_1,r\mathbf{v}_1)\} \tag{11}$$

leads now to a contribution of the kinetic energy to the acceptance probability of Eq. (9), which is given by

$$\Delta K(r) = (\beta_2 - \beta_1 r^{-2})[E_{kin}(\mathbf{v}_2) - r^2E_{kin}(\mathbf{v}_1)], \tag{12}$$

where we have used the property $E_{kin}(r\mathbf{v})=r^2E_{kin}(\mathbf{v})$. In order to optimize the acceptance of replica exchange moves one can adjust the scale r in such a way that the kinetic energy contribution to the Metropolis term vanishes. The condition $\Delta K(r)=0$ has two solutions. Following Ref. [9] one can choose

$$r = r_\beta \equiv \sqrt{\frac{\beta_1}{\beta_2}} = \sqrt{\frac{T_2}{T_1}}. \tag{13}$$

This choice of r leads to $\beta_2-\beta_1 r^{-2}=0$ in Eq. (12), but obviously this is not the only possibility. An alternative is to adjust r in a way that $E_{kin}(\mathbf{v}_2)-r^2E_{kin}(\mathbf{v}_1)=0$. This can be realized by setting

$$r = r_E \equiv \sqrt{E_{kin}(\mathbf{v}_2)/E_{kin}(\mathbf{v}_1)} = \sqrt{\hat{T}_2/\hat{T}_1}. \tag{14}$$

Here, $\hat{T}=2k_BE_{kin}$ is the instantaneous temperature of the velocity field, as opposed to the thermostat temperature T. Obviously, in the thermodynamic limit $\hat{T}\to T$, and both scalings become identical. However, in finite systems $\hat{T}$ fluctuates around T and, in general, one has $r_E \neq r_\beta$.

As both scalings preserve detailed balance [14], they lead to equally valid but different replica exchange moves. Both scalings are optimal in the sense that they render acceptance rates independent of fluctuations in the velocity field. Current implementations usually employ only a single variant. Choosing randomly among both moves can increase mixing in replica exchange, albeit without further increasing acceptance rates.

Since $\Delta K(r)\to-\infty$ for $r\to 0$ as well as for $r\to\infty$ and $\Delta K(r_\beta)=0=\Delta K(r_E)$, there exists a regime of r where $\Delta K(r)>0$. It is therefore tempting to maximize $\Delta K(r)$ as this will increase the acceptance probability of an exchange move. From the condition

$$\frac{d}{dr}\Delta K(r) = \beta_2 E_{kin}(\mathbf{v}_1) - \frac{\beta_1 E_{kin}(\mathbf{v}_2)}{r^4} \equiv 0, \tag{15}$$

one finds

$$r_{opt} = \sqrt[4]{\frac{\beta_1 E_{kin}(\mathbf{v}_2)}{\beta_2 E_{kin}(\mathbf{v}_1)}} = \sqrt{r_\beta r_E}, \tag{16}$$

which leads to a positive maximal contribution of the kinetic energy given by

$$\Delta K(r_{opt}) = [\sqrt{\beta_1 E_{kin}(\mathbf{v}_1)} - \sqrt{\beta_2 E_{kin}(\mathbf{v}_2)}]^2. \tag{17}$$

Hence, with such a rescaling of velocities the fluctuations of the instantaneous temperatures of the velocity fields can be utilized to *increase* the acceptance probability of exchange moves, allowing for larger jumps in potential energy. Note, however, that the rescaling according to Eq. (16) preserves detailed balance only in the thermodynamic limit—i.e., where $\hat{T}\to T$ holds and $\Delta K(r_{opt})\equiv 0$ anyhow. For finite systems and during equilibration—i.e., where $\hat{T}$ deviates from T and $\Delta K(r_{opt})>0$—detailed balance is violated, albeit to a lesser degree the larger the system is. For this reason, the scaling of Eq. (16) has to be used with care.

The above theoretical framework allows us to introduce also an interesting variant of replica exchange MD. Here the stochastic sampler is just the regular MD simulation without a heat bath. Provided the dynamics is ergodic, a microcanonical MD simulation can be viewed as a constant sampling on the energy surface, $E=E_{kin}+E_{pot}=\text{const}$. The corresponding weight function is $w(S)=1/|\Omega(E)|$, where $\Omega(E)$ denotes

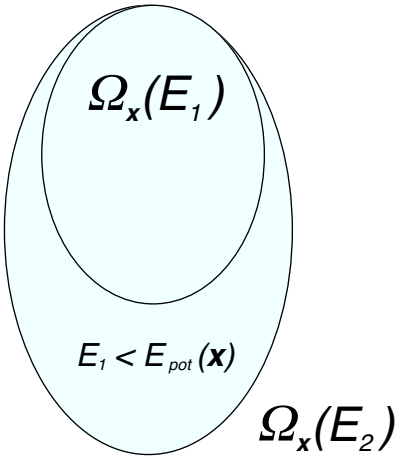


FIG. 2. Sketch of the coordinate part of the phase space for constant-energy replica exchange simulations. Exchange moves are possible in the overlap region.

the phase space of the hypersurface of constant energy at E. It is well known that microcanonical MD exhibits slow equilibration and independently sampled trajectories should be combined to ensure better statistics. However, this approach can be extended readily into a replica exchange scheme for speeding up equilibration on several energy surfaces together. We assume $E_1 < E_2$ in the following. The move set is a generalization of Eq. (11),

$$\{(\mathbf{x}_1,\mathbf{v}_1),(\mathbf{x}_2,\mathbf{v}_2)\} \to \{(\mathbf{x}_2,r_2\mathbf{v}_2),(\mathbf{x}_1,r_1\mathbf{v}_1)\}, \tag{18}$$

involving two different rescaling factors r_1 and r_2:

$$r_{1,2} = \sqrt{\frac{E_{2,1} - E_{pot}(\mathbf{x}_{1,2})}{E_{1,2} - E_{pot}(\mathbf{x}_{1,2})}}. \tag{19}$$

Such moves are possible for $E_{pot}(\mathbf{x}_2) < E_1$ (see Fig. 2); $E_{pot}(\mathbf{x}_1) < E_2$ automatically holds. We note that this restriction does not violate detailed balance. Furthermore, the combination with the regular MD simulation ensures ergodicity.

A fascinating aspect of this scheme is that the acceptance probability for an allowed move is always 1, since both weight functions are constant. No other scenario allows for such a high acceptance rate. Using reweighting techniques [17], canonical properties can be obtained from simulations at several different energy values. Moreover, constant-energy surface simulations may be of interest in their own right [10,11], e.g., for comparison with recent molecular beam experiments [18].

Practical acceptance probabilities will be somewhat smaller than 1, $\langle p_M \rangle < 1$, due to the region of forbidden moves. Using the distribution of the potential energy on energy shell E_i, $P_{E_i}(E_{pot})$, the average acceptance probability is given by

$$\langle p_M \rangle = \int_{E_{pot}^{(min)}}^{E_1} dE_{pot} P_{E_2}(E_{pot}). \tag{20}$$

For classical trajectories of total energy E equipartition of average kinetic and potential energy usually holds,

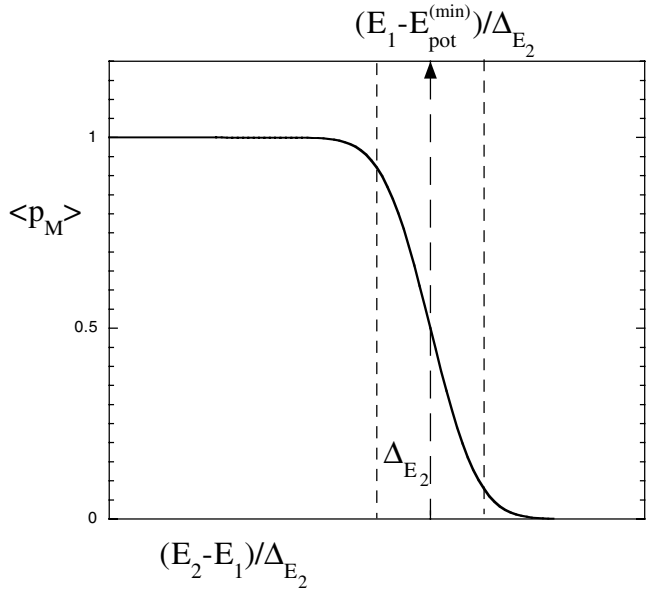


FIG. 3. Dependence of the average acceptance probability $\langle p_M \rangle$ on the distance of the energy shells [see Eq. (23)] in microcanonical replica exchange simulations.

$$\langle E_{kin} \rangle_E = \langle E_{pot} - E_{pot}^{(min)} \rangle_E = (E - E_{pot}^{(min)})/2, \tag{21}$$

with $E_{pot}^{(min)}$ the energy of the lowest-energy configuration. Assuming a Gaussian distribution for the potential energy,

$$P_{E_i}(E_{pot}) \propto \exp[-(E_{pot} - \langle E_{pot} \rangle_{E_i})^2/\Delta_{E_i}], \tag{22}$$

we obtain the dependence of the average acceptance probability $\langle p_M \rangle$ on the energy difference $E_2 - E_1$,

$$\langle p_M \rangle \approx \frac{1}{2} \operatorname{erfc}\left(\frac{E_2 - E_1}{\Delta_{E_2}} - \frac{E_1 - E_{pot}^{(min)}}{\Delta_{E_2}} \right), \tag{23}$$

shown in Fig. 3. In particular, $\langle p_M \rangle$ will be at least one-half if the average potential energy at E_2 is smaller than E_1, $\langle E_{pot} \rangle_{E_2} < E_1$. This criterion is equivalent to the energy difference being at most equal to twice the average kinetic energy at E_1, $E_2 - E_1 < E_1 - E_{pot}^{(min)} = 2\langle E_{kin} \rangle_{E_1}$. Figure 3 shows that $\langle p_M \rangle$ will rapidly approach 1 upon decreasing energy difference.

This last example demonstrates the wide applicability of generalized replica exchange move sets. It also demonstrates the striking advantage of replica exchange over earlier approaches like simulated tempering [19]. In the latter, additional parameters reflecting free-energy differences are of utmost importance. Their determination is tedious, and approximations [20] are useful only in certain limiting cases. In replica exchange such normalization constants simply drop out due to the form of the acceptance probability, Eq. (3). Constant-energy surface simulations as sketched above could be approached in simulated tempering only with a solid knowledge of the phase-space ratios, $|\Omega(E_2)|/|\Omega(E_1)|$.

In summary, we have sketched the statistical physics framework of applying the replica exchange technique to MD simulations. Generalized move sets—in particular, ap-

propriate rescaling of the velocity fields—allow optimization of the acceptance probability as well as new approaches. Together with an optimization of the temperature spacing to increase replica flow [15,16] optimized acceptance probabilities will lead to shorter simulation times in canonical replica exchange MD simulations. Microcanonical replica exchange MD simulations are intrinsically optimized and will provide new insights.

It is a pleasure to thank S. Hoefinger and P. Grassberger for discussions. This research was supported by NSF Grant No. CHE-0313618.

[1] J. Skolnick and A. Kolinski, Comput. Sci. Eng. **3**(9/10), 40 (2001).
[2] K. Binder and W. Kob, *Glassy Materials and Disordered Solids: An Introduction to their Statistical Mechanics* (World Scientific, Singapore, 2005).
[3] *Proceedings of the Erice workshop on Complexity, Metastability and Nonextensivity*, edited by C. Beck, G. Benedek, A. Rapisarda, and C. Tsallis (World Scientific, Singapore, 2005).
[4] C. J. Geyer and A. Thompson, J. Am. Stat. Assoc. **90**, 909 (1995).
[5] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).
[6] U. H. E. Hansmann, Chem. Phys. Lett. **281**, 140 (1997).
[7] There exist approaches to allow exchange between all samplers; see, e.g., P. Brenner, C. R. Sweet, D. VonHandorf, and J. A. Izaguirre, J. Chem. Phys. **126**, 074103 (2007), and references therein; our considerations can be applied also to such approaches.
[8] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[9] Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **314**, 141 (1999).
[10] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, Oxford, 1989).
[11] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, San Diego, 2002).
[12] For the limitations of various approaches using the velocity field as heat bath see Ref. [11], Chap. 6.1.
[13] S. Nose, J. Chem. Phys. **81**, 511 (1984).
[14] Any state-independent scale r preserves detailed balance (DB). In addition, there exist two classes of state-dependent scalings $r=r(\mathbf{v}_1,\mathbf{v}_2)$ that also preserve DB: One is $r=r_E$, Eq. (14), and the other is any function of a product of the kinetic energies, $r=f[E_{kin}(\mathbf{v}_1)E_{kin}(\mathbf{v}_2)]$. Of all scalings that are compatible with DB, only $r=r_\beta$ and $r=r_E$ are distinguished in that they lead to $\Delta K(r)\equiv 0$.
[15] S. Trebst, M. Troyer, and U. H. E. Hansmann, J. Chem. Phys. **124**, 174903 (2006).
[16] W. Nadler and U. H. E. Hansmann, Phys. Rev. E **75**, 026109 (2007).
[17] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988); **63**, 1658(E) (1989), and references given in the erratum.
[18] M. Kohtani, T. C. Jones, J. E. Schneider, and M. F. Jarrold, J. Am. Chem. Soc. **126**, 7420 (2004).
[19] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).
[20] S. Park and V. S. Pande, Phys. Rev. E **76**, 016703 (2007).