

Rigorous classical-mechanical derivation of a multiple-copy algorithm for sampling statistical mechanical ensembles

Christopher Adam Hixson and Ralph A. Wheeler

Department of Chemistry and Biochemistry, Room 208, University of Oklahoma, 620 Parrington Oval, Norman, Oklahoma 73019

(Received 26 March 2001; published 10 July 2001)

We derive a rigorous, multiple-copy simulation algorithm that is formally equivalent to conventional classical molecular dynamics for an ensemble of systems, but may be used for rapid geometry optimizations. The derivation is accomplished by starting from an ensemble of copies of the entire system and applying a point coordinate transformation to a large subsystem defined as the bath. After the transformation, each atom of the bath is described by one “major” set of coordinates located at the average position of the ensemble of equivalent atoms and a set of “minor” coordinates that when combined with the “major” coordinates represent exact dynamics. Neglecting the “minor” set of coordinates results in a Hamiltonian and a probability density equivalent to those used in existing multiple-copy methods. Neglecting Hamilton’s equations of motion for the minor variables gives the equations of motion for locally enhanced sampling. Numerical tests indicate that the algorithm can recover exact molecular dynamics of the ensemble, conventional multiple-copy dynamics, or results of intermediate accuracy. Thus, the algorithm provides a rigorous basis for multiple-copy dynamics, resolves many of the uncertainties associated with their current implementations, and offers the potential for calculating ensemble average properties in conjunction with finding a system’s global minimum energy geometry.

DOI: 10.1103/PhysRevE.64.026701

PACS number(s): 45.10.-b, 02.70.Ns, 45.20.Jj, 87.15.Aa

I. INTRODUCTION

More than ten years ago, Elber and Karplus presented a multiple-copy molecular dynamics method designed to accelerate a simulation’s convergence to a global energy minimum [1]. The method, called locally enhanced sampling (LES), is based on creating a set of non-interacting copies of a small subsystem of primary interest and allowing a larger subsystem, the “bath,” to interact with each copy of the subsystem. The force each copied atom experiences from the bath is the total force the corresponding real atom would experience. The bath atoms, on the other hand, experience the average of the forces due to the copied atoms. As a result, energy barriers that copies of the subsystem must overcome to avoid being trapped in local energy minima are decreased compared to those in a conventional molecular dynamics simulation. LES and related mean field methods, have been used successfully in a variety of optimization problems [1–17], but they suffer limitations common to other *ad hoc* geometry optimization methods [18–26]: since the underlying energy surface and/or its sampling is modified, the methods do not generate trajectories that correspond to any of the familiar statistical mechanical ensembles. So, phase-space averaging over these trajectories is not useful in a statistical mechanical sense and any information obtained in this manner (free energy differences, radial distribution functions, or even temperatures) must be used with caution. It should come as no surprise, then, that data acquired from LES has been found to violate fundamental principles, including the equipartition of energy theorem [27,28]. It has been claimed [27–29] that this manifests itself as the “temperature disparity problem,” which is a failure of the subsystem and bath temperatures to reach the same equilibrium value. On a separate issue, Stultz and Karplus have provided a proof that

minima located using LES are not necessarily minima on the original energy surface [30].

Although optimization methods based on mean field theory show great practical utility [1–17,31–33], current multiple-copy implementations are clearly flawed. The current contribution was inspired by LES, but our approach is intended to improve phase-space sampling by using a rigorously derived method, so our approximations are well known and may be controlled. By controlling the approximations, we hope to generate trajectories that more closely approximate trajectories expected of systems belonging to one of the well-studied statistical mechanical ensembles. Our starting point is fundamentally different from that of Elber and Karplus, but it is nonetheless instructive to reiterate briefly their rationale for LES. Next we describe known drawbacks of LES and attempts to remedy them. Finally, we derive our generalization of multiple-copy methods, present a numerical test, and relate our algorithm to existing multiple-copy methods. Calculation of thermodynamic quantities will be reported separately.

II. CONVENTIONAL MULTIPLE-COPY MOLECULAR DYNAMICS

As it was originally presented [1], LES flows from the assumption that the classical phase-space density can be written as

$$\rho(\bar{X}, t) = \rho_s(\bar{X}_s, t) \rho_b(\bar{X}_b, t), \quad (1)$$

where ρ_s is the density of the subsystem to be copied, ρ_b is the density of the bath, and \bar{X} is the vector that indicates the system’s location in phase space. It is also assumed that the bath’s density can be written as a single delta function:

$$\rho_b(\bar{X}_b, t) = \delta(\bar{X}_b(t)), \quad (2)$$

while the copied subsystem's density can be taken to be a "swarm" of delta functions, such as

$$\rho_s(\bar{X}_s, t) = \sum_{k=1}^C w_{s,k} \delta(\bar{X}_{s,k}(t)), \quad (3)$$

where these represent the positions in phase space of the various copies, and $w_{s,k}$ is a weighting function. In LES, the $w_{s,k}$ are taken to be $1/C$, where C is the number of copies. Requiring that this form of the phase-space density satisfies Hamilton's equations of motion for ensemble averages of the individual particles and momenta (derived from the Liouville equation [34]) gives analogous differential equations describing motions of the bath atoms and the copied system atoms:

$$\dot{q}_{i,k} = \frac{\partial H_k}{\partial p_{i,k}}, \quad (4a)$$

$$\dot{p}_{i,k} = -\frac{\partial H_k}{\partial q_{i,k}}; \quad (4b)$$

$$\dot{Q}_i = \sum_{k=1}^C w_k \frac{\partial H_k}{\partial P_i}, \quad (4c)$$

$$\dot{P}_i = -\sum_{k=1}^C w_k \frac{\partial H_k}{\partial Q_i}. \quad (4d)$$

Here, lower case variables refer to the copied subsystem and uppercase variables refer to the "bath." The i refers to the coordinate index, while k indicates the copy. The chosen form for the phase-space density is never rigorously justified, but rather is rationalized by analogy to the time-dependent Hartree (TDH) approximation for quantum mechanical wave functions. This analogy was previously used to obtain a similar, rigorous method intended for small molecules [35], after taking the quantum expression to the classical limit. The original method is rigorous because the entire (small) system is replicated and "trajectory bundles" are simulated, whereas LES replicates only a small part of the system. Although the quantum mechanical analogy is very clever and insightful, it provides limited prospects for improving approximate multiple-copy simulation methods and has left many workers wondering exactly how the approximation alters the dynamics.

To understand and resolve the limitations and uncertainties inherent in LES, subsequent workers have tried to put multiple-copy methods on a more stable foundation. Most notably, Zheng and Zheng [29] claim a derivation of LES by starting with copies of the entire system. Then, they perform a unitary transformation of coordinates for particles of the bath and integrate out of the phase-space density all of the transformed variables except for those corresponding to the mean of the bath's coordinates. The transformation employed was of the form

$$Q_i = \frac{1}{n} \left(\sum_{k=1}^C Q_{i,k} \right), \quad (5a)$$

$$P_i = \frac{1}{n} \left(\sum_{k=1}^C P_{i,k} \right), \quad (5b)$$

$$Q'_{i,l} = \frac{1}{n} \left(Q_{i,1} + \sum_{k=2}^C c_{l,k} Q_{i,k} \right), \quad (5c)$$

$$P'_{i,l} = \frac{1}{n} \left(P_{i,1} + \sum_{k=2}^C c_{l,k} P_{i,k} \right), \quad (5d)$$

where we refer to $\{Q_i\}_{i=1}^N$ and $\{P_i\}_{i=1}^N$ as the "major" variables, $\{Q'_{i,l}\}_{i=1,l=2}^{N,C}$ and $\{P'_{i,l}\}_{i=1,l=2}^{N,C}$ as the "minor" variables, and $\{c_{i,j}\}$ are the transform coefficients. The major variables correspond to the average position of the bath coordinates for all of the copies, and the minor variables can be viewed as an orthogonal set of vectors that describe the fine details of the dynamics. There are, of course, constraints that define the transformation. Not only must the coefficients generate an orthogonal transformation, it is also convenient to normalize them such that

$$1 + \sum_{k=2}^C c_{n,k} c_{m,k} = C \delta_{m,n}$$

with the additional constraint that

$$1 + \sum_{k=2}^C c_{n,k} = 0.$$

In their work, Zheng and Zheng used a probability density appropriate for the canonical ensemble and only considered the special case of a harmonic potential energy function. After transforming and integrating out the minor variables, their result seemed to imply that by choosing the value of n used in the transformation above, the temperature disparity problem could be solved. The choice of $n = \sqrt{C}$ makes the transformation canonical, which when applied to the arguments in [27] and [28], the authors felt, would correct the equipartition problem for the harmonic potential considered.

To investigate the general temperature disparity problem, an originator of LES showed that LES conserves the following Hamiltonian [28]:

$$\begin{aligned} H_{LES} = & \sum_{i=1}^N \frac{p_i^2}{2m_i} + \sum_{i>j=1}^N V(Q_i, Q_j) \\ & + \frac{1}{C} \sum_{k=1}^C \left(\sum_{i=1}^S \frac{p_{i,k}^2}{2m_i} + \sum_{i>j=1}^S V(q_{i,k}, q_{j,k}) \right. \\ & \left. + \sum_{i=1}^S \sum_{j=1}^N V(q_{i,k}, Q_j) \right), \end{aligned} \quad (6)$$

where S is the number of copied particles, C is the number of times the particles are copied, and N is the number of uncopied "bath" particles. Lowercase variables refer to the copied particles, while uppercase variables indicate the "bath." Dynamical information is obtained by integrating the equations

$$\dot{q}_{i,k} = C \frac{\partial H_{LES}}{\partial p_{i,k}}, \quad (7a)$$

$$\dot{p}_{i,k} = -C \frac{\partial H_{LES}}{\partial q_{i,k}}; \quad (7b)$$

$$\dot{Q}_i = \frac{\partial H_{LES}}{\partial P_i}, \quad (7c)$$

$$\dot{P}_i = -\frac{\partial H_{LES}}{\partial Q_i}. \quad (7d)$$

This is an approximation to the dynamics of C systems.

Straub and Karplus, who originally introduced the Lagrangian generator of this Hamiltonian, propose energy scaling as a solution for the temperature disparity problem [27]. Ulitsky and Elber derived the Hamiltonian given above [28] and noted that the TDH approximation does not follow the classical virial theorem and therefore violates energy equipartitioning, and claimed that this fault is the source of some of the more apparent limitations of LES. To remedy this, Ulitsky and Elber appealed to the Boltzmann equation of transport theory and created an algorithm that monitored ‘‘collisions’’ between atoms of copied particles and the bath, in analogy to the collision integral. Then, they modified their dynamics to treat the collision exactly, instead of using the LES approximation. With this ‘‘collisional’’ LES (cLES) scheme, they achieved improved behavior in the dynamics of systems studied [28,31].

III. RIGOROUS ALGORITHM FOR MULTIPLE-COPY DYNAMICS

Inspired by the insightful approach of Zheng and Zheng [29], we decided to take an exact expression for an ensemble of C independent copies of a system, and transform it to obtain an exact multiple-copy dynamics for systems with a more general potential energy function than the harmonic form considered by Zheng and Zheng. Using a transform similar to the one given above makes this possible. The key is to realize that the reverse transform has the form

$$Q_{i,1} = \frac{C}{n} \left(Q_i + \sum_{k=2}^C Q'_{i,k} \right), \quad (8a)$$

$$P_{i,1} = \frac{C}{n} \left(P_i + \sum_{k=2}^C P'_{i,k} \right), \quad (8b)$$

$$Q_{i,k} = \frac{C}{n} \left(Q_i + \sum_{l=2}^C c_{l,k} Q'_{i,l} \right), \quad (8c)$$

$$P_{i,k} = \frac{C}{n} \left(P_i + \sum_{l=2}^C c_{l,k} P'_{i,l} \right), \quad (8d)$$

which can then be applied to the exact Hamiltonian of C noninteracting copies of a system. This can be done by first

writing the Hamiltonian of a collection of C equivalent copies of the same system assuming a two-body form for the potential energy:

$$H = \sum_{k=1}^C \left(\sum_{i=1}^S \frac{p_{i,k}^2}{2m_i} + \sum_{i>j=1}^S V(q_{i,k}, q_{j,k}) + \sum_{i=1}^S \sum_{j=1}^N V(q_{i,k}, Q_{j,k}) + \sum_{i=1}^N \frac{P_{i,k}^2}{2M_i} + \sum_{i>j=1}^N V(Q_{i,k}, Q_{j,k}) \right). \quad (9)$$

Here the first two terms are the energy due to the copied subsystem, the middle term is the interaction energy between the copies and the bath, and the final two terms are the energy due to the bath. Applying the transform yields

$$\begin{aligned} \tilde{H} = & \sum_{i=1}^S \sum_{k=1}^C \frac{p_{i,k}^2}{2m_i} + \sum_{i>j=1}^S \sum_{k=1}^C V(q_{i,k}, q_{j,k}) + \sum_{j=1}^N \frac{n^2 P_j^2}{2M_j C} \\ & + \sum_{i>j=1}^N \sum_{k=1}^C V \left(\frac{C}{n} \left(Q_i + \sum_{m=2}^C c_{m,k} Q'_{i,m} \right) \right), \\ & \frac{C}{n} \left(Q_j + \sum_{m=2}^C c_{m,k} Q'_{j,m} \right) \Bigg) + \sum_{j=1}^C \sum_{k=2}^C \frac{n^2 P'_{j,k}}{2M_j C} \\ & + \sum_{i=1}^S \sum_{j=1}^N \sum_{k=1}^C V \left(q_{i,k}, \frac{C}{n} \left(Q_j + \sum_{m=2}^C c_{m,k} Q'_{j,m} \right) \right). \end{aligned} \quad (10)$$

Here, again, the first two terms are the energy of the copied particles. The transform has split the bath energy into three terms in Eq. (10): the kinetic energy of the major variables, the potential energy due to the bath, and the kinetic energy due to the minor variables. The last term in Eq. (10) is the interaction energy between the copies and the bath.

Obviously, the original Hamiltonian can be used to generate equations of motion, which can be integrated to generate the desired trajectories. The transformed Hamiltonian's equations of motions, though, are more interesting. The untransformed, ‘‘copied’’ part is just the same as in the untransformed representation, but the ‘‘bath’’ part is more complex. For example, the time derivative,

$$\dot{Q}_i = \frac{1}{n} \left(\sum_{k=1}^C \dot{Q}_{i,k} \right),$$

can be rewritten as

$$\dot{Q}_i = \frac{1}{n} \left(\sum_{k=1}^C \frac{\partial \tilde{H}}{\partial P_i} \frac{\partial P_i}{\partial P_{i,k}} \right),$$

which can finally be reduced to

$$\dot{Q}_i = \frac{C}{n^2} \frac{\partial \tilde{H}}{\partial P_i}. \quad (11)$$

Given all of this, the entire set of equations of motion can be rewritten as

$$\dot{q}_{i,k} = \frac{\partial \tilde{H}}{\partial p_{i,k}}, \quad (12a)$$

$$\dot{p}_{i,k} = -\frac{\partial \tilde{H}}{\partial q_{i,k}}; \quad (12b)$$

$$\dot{Q}_i = \frac{C}{n^2} \frac{\partial \tilde{H}}{\partial P_i}, \quad (12c)$$

$$\dot{P}_i = -\frac{C}{n^2} \frac{\partial \tilde{H}}{\partial Q_i}; \quad (12d)$$

$$\dot{P}'_{i,k} = -\frac{C}{n^2} \frac{\partial \tilde{H}}{\partial Q'_{i,k}}, \quad (12e)$$

$$\dot{Q}'_{i,k} = \frac{C}{n^2} \frac{\partial \tilde{H}}{\partial P'_{i,k}}. \quad (12f)$$

These equations are the exact equations of motion in the transformed representation. It should be noted that if $n = \sqrt{C}$, we have a canonical transformation, as Zheng and Zheng accomplished for harmonic potentials. But if $n = C$, the point transform generated yields equations of motion that are very similar to the LES equations. (Other choices of n lead to other, equivalent representations of the dynamics, and may ultimately prove important.) Now, if all of the bath particles start from the same initial positions with the same velocities in all of the copies and all of the minor variables are ignored (assumed to vanish), the LES equations of motion are recovered. We claim that all of the error present in this sort of multiple-copy method comes from this holonomic constraint and that previously noted faults flow from this idea. Numerical simulations presented in Fig. 1 support this view for the temperature disparity problem. The flattest curve in Fig. 1 shows the temperature of an argon bath as a function of time and the curve that oscillates around it is the temperature of a single argon atom calculated using exact dynamics. The top, bold curve is the temperature of a single argon atom, represented by four copies, in the LES approximation. Clearly, the LES particle's temperature does not relax to equilibrium as it should, and its fluctuations are much larger than in exact dynamics. Our algorithm can be used to generate each curve in Fig. 1, including those intermediate between LES and exact dynamics. These curves were generated using a scheme practically similar to cLES, but with a fundamental difference in implementation and basis. In cLES, "collisions" were detected by physical proximity between a copied particle and a bath particle, whereas the correction presented here calculates the "minor" variables' forces (the \dot{P}' 's from above), which are neglected in both LES and cLES. Our method still ignores these forces when they are sufficiently small, but when the force exerted on the minor variable is greater than some threshold, the particle is

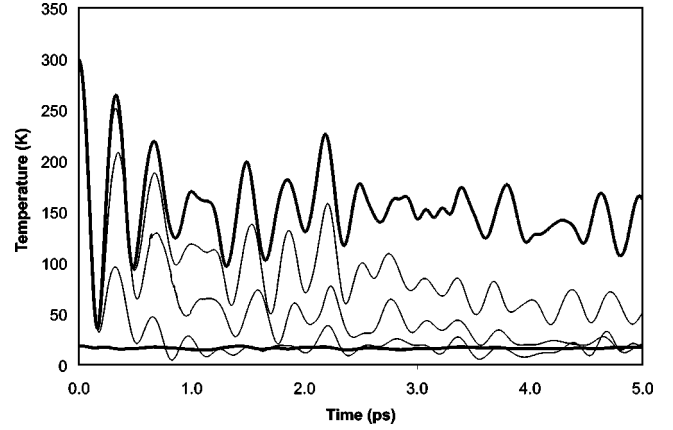


FIG. 1. The flattest curve is the temperature as a function of time for a bath of 64 Ar atoms calculated using molecular dynamics. The curve that oscillates around it is the temperature of a single Ar atom relaxing to equilibrium calculated using exact molecular dynamics. The bold curve is the temperature of one Ar calculated using the LES approximation. The curves between the exact and LES curves were calculated using the algorithm presented.

then removed from the bath and treated exactly, as if a copied particle. By adjusting the threshold force, the dynamics can be scaled from purely LES to exact dynamics, in principle visiting all points in between. This allows the computational advantage of LES to be largely maintained, while providing a rigorous procedure that leads to improved dynamics.

IV. COMPARISONS WITH CONVENTIONAL MULTIPLE-COPY DYNAMICS

As most of the work produced to explain conventional multiple-copy dynamics depends on a phase-space density approach [1,2,27–29] (often involving the Liouville equation), further comparisons of our work with previously published methods would be helped by casting our work in such a formalism. First, we consider a formal representation of the density that satisfies the Liouville equation [34],

$$\rho(\bar{X}, t) = e^{-\hat{L}t} \rho(\bar{X}, 0) \quad (13)$$

where \hat{L} is a Liouville operator,

$$\hat{L} = \sum_{i=1}^S \sum_{k=1}^C \left(\frac{\partial H}{\partial p_{i,k}} \frac{\partial}{\partial q_{i,k}} - \frac{\partial H}{\partial q_{i,k}} \frac{\partial}{\partial p_{i,k}} \right) + \sum_{i=1}^N \sum_{k=1}^C \left(\frac{\partial H}{\partial P_{i,k}} \frac{\partial}{\partial Q_{i,k}} - \frac{\partial H}{\partial Q_{i,k}} \frac{\partial}{\partial P_{i,k}} \right), \quad (14)$$

and $\rho(\bar{X}, 0)$ is the density at time zero:

$$\rho(\bar{X}, 0) = \left(\prod_{i=1}^S \prod_{k=1}^C \delta(q_{i,k} - q_{i,k,0}) \delta(p_{i,k} - p_{i,k,0}) \right) \times \left(\prod_{i=1}^N \prod_{k=1}^C \delta(Q_{i,k} - Q_{i,k,0}) \delta(P_{i,k} - P_{i,k,0}) \right). \quad (15)$$

Next, we apply the same point coordinate transformation as before to both the Liouville operator and the initial probability density. This generates the transformed Liouvillian

$$\hat{L} = \sum_{i=1}^S \sum_{k=1}^C \left(\frac{\partial \tilde{H}}{\partial p_{i,k}} \frac{\partial}{\partial q_{i,k}} - \frac{\partial \tilde{H}}{\partial q_{i,k}} \frac{\partial}{\partial p_{i,k}} \right) + \frac{C}{n^2} \left[\sum_{i=1}^N \left(\frac{\partial \tilde{H}}{\partial P_i} \frac{\partial}{\partial Q_i} - \frac{\partial \tilde{H}}{\partial Q_i} \frac{\partial}{\partial P_i} \right) + \sum_{i=1}^N \sum_{k=2}^C \left(\frac{\partial \tilde{H}}{\partial P'_{i,k}} \frac{\partial}{\partial Q'_{i,k}} - \frac{\partial \tilde{H}}{\partial Q'_{i,k}} \frac{\partial}{\partial P'_{i,k}} \right) \right]. \quad (16)$$

After applying the Jacobian of the transformation, $J(\bar{X}, \bar{X})$, and assuming that the bath's initial positions and velocities are the same in all of the copies, the transform also generates the density [36],

$$\begin{aligned} \tilde{\rho}(\bar{X}, 0) &= \frac{1}{|J(\bar{X}, \bar{X})|} \left(\prod_{i=1}^S \prod_{k=1}^C \delta(q_{i,k} - q_{i,k,0}) \delta(p_{i,k} - p_{i,k,0}) \right) \\ &\times \left(\prod_{i=1}^N \delta(Q_i - Q_{i,0}) \delta(P_i - P_{i,0}) \right) \\ &\times \left(\prod_{k=2}^C \delta(Q'_{i,k}) \delta(P'_{i,k}) \right). \end{aligned} \quad (17)$$

Thus the exact time-dependent phase-space density in the transformed representation, $\tilde{\rho}(\bar{X}, t)$, is formed by using the transformed Liouvillian and transformed initial densities in the above formal expression. One can then follow the example provided in Zheng and Zheng's work to arrive at a reduced density, $\tilde{\rho}'(\bar{X}', t)$, that can be used to generate LES-type equations of motion. This is simply done by integrating over the minor variables

$$\tilde{\rho}'(\bar{X}', t) = \int d\gamma' |J(\bar{X}, \bar{X})| e^{\hat{L}t} \tilde{\rho}(\bar{X}, 0), \quad (18)$$

where $d\gamma'$ is the volume element in the space that the minor variables occupy. Not only does this integration directly generate a density that can be used to obtain the LES equations of motion, but it also serves to enforce the holonomic constraints that were found above to result in LES.

The approach presented to generate a multiple-copy method also sheds some light onto the violation of energy equipartitioning [27] or equivalently, the incorrect virial given by LES [28]. One of the well known results of classical mechanics is that the time (or assuming ergodicity, the ensemble) average of the kinetic energy can be expressed in the following way [37]:

$$\bar{T} = -\frac{1}{2} \sum_i \langle q_i F_i \rangle, \quad (19)$$

where $\langle A \rangle$ is ensemble average of quantity A , and the summation is taken over all the degrees of freedom in the system. In the canonical ensemble, this produces the familiar

result that $\bar{T} = N/2\beta$, where N is the number of degrees of freedom in the system, β is $1/kT$, and it is assumed that the potential diverges on the surface of the volume of integration. This assumption corresponds physically to considering a bound system. In this special case of a bound system, Ulitsky and Elber showed that the virial theorem generates a strange result, $\bar{T} = (N + CS)/2\beta$, when applied in [28]. From this expression, it appears that the temperature of the copied subsystem is C times hotter than the bath's. Our approach confirms this result, but demonstrates that this too can be traced to a neglect of the minor variables.

In the canonical ensemble, evaluating $\langle q_i F_i \rangle$ for a bound system entails solving the integral

$$\langle q_i F_i \rangle = -\frac{1}{Z} \int d\Gamma dq_i q_i \frac{\partial H}{\partial q_i} e^{-\beta H(q_i, \Gamma)}. \quad (20)$$

Here, Z is the partition function and q_i can be any of the coordinates described in this paper, be it major, minor, or untransformed. The integral is evaluated most easily by integrating by parts and its exact value is $-1/\beta$. Therefore, the virial theorem generates an average kinetic energy for an ensemble of C systems as simply

$$\begin{aligned} \bar{T} &= -\frac{1}{2} \sum_{i=1}^S \sum_{k=1}^C \left\langle q_{i,k} \frac{\partial H}{\partial q_{i,k}} \right\rangle - \frac{1}{2} \sum_{i=1}^N \sum_{k=1}^C \left\langle Q_{i,k} \frac{\partial H}{\partial Q_{i,k}} \right\rangle \\ &= \frac{(SC + NC)}{2\beta}. \end{aligned} \quad (21)$$

After making our transformation,

$$\begin{aligned} \left\langle Q_{i,k} \frac{\partial H}{\partial Q_{i,k}} \right\rangle &= \left\langle \frac{C}{n} \left(Q_i + \sum_{l=2}^C c_{l,k} Q'_{i,l} \right) \right. \\ &\times \left. \frac{C}{n^2} \left(\frac{\partial H}{\partial Q_i} + \sum_{m=2}^C c_{m,k} \frac{\partial H}{\partial Q'_{i,m}} \right) \right\rangle, \end{aligned} \quad (22)$$

and using the result in the virial theorem, the exact result, $\bar{T} = (SC + NC)/2\beta$, is generated. However, if the minor variables are ignored, one obtains the result $\bar{T} = (N + CS)/2\beta$. So, the problem with the LES virial is simply a counting problem and the average kinetic energy takes the Ulitsky-Elber form because of the reduced number of degrees of freedom found in LES.

V. CONCLUSIONS

Starting from an ensemble of identical systems and applying a point transformation to the coordinates of a large number of "bath" particles generates an algorithm for efficiently replicating the dynamics of the ensemble. The transformation gives a description of the bath in terms of "major" variables located at the average phase-space position of equivalent atoms and a set of "minor" variables describing the finer details of the bath dynamics. Numerical tests show that the algorithm can recover exact dynamics or give dynamics identical to conventional multiple-copy dynamics [1], if the

minor variables are neglected. If the minor variables are included in the dynamics only when the corresponding forces exceed a chosen threshold, results of intermediate accuracy are obtained. Varying the threshold force controls the accuracy of the calculation and the computer time required.

Applying the point transformation to an ensemble of identical systems also opens a new perspective on conventional multiple-copy dynamics. First, neglecting Hamilton's equations of motion for the minor variables of the bath gives the equations of motion that define LES [1]. Second, applying the same point transformation to the Liouville equation and probability density, followed by enforcing the holonomic constraint that the minor variables vanish, gives a probability density similar to that assumed for LES. Third, neglecting the minor variables gives the same incorrect virial as LES, but including the minor variables yields the correct virial for the ensemble. Finally, the success of collisional LES is understandable, as cLES includes the minor variable dynamics in an empirical way.

Since mean field theories such as LES have proven extremely useful for locating global minima on complex free energy surfaces [1–17], work is currently underway to de-

velop the algorithm described here into a similar tool. In addition, we are investigating the limitations of the algorithm for calculating ensemble average properties accurately and quickly.

ACKNOWLEDGMENTS

This work was funded by the Chemical Sciences, Geosciences and Bioscience Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy through Grant No. DE-FG03-01ER15164 and by the Oklahoma Center for the Advancement of Science and Technology through OCAST Grant No. HR01-148. We are grateful to the University of Oklahoma's Graduate College for supporting C.A.H. R.A.W. thanks Peter Kollman's research group for their hospitality during his sabbatical, when this project was conceived. We are especially grateful to Professor Carlos L. Simmerling for helpful discussions concerning the limitations of LES. We also thank Professor Kieran Mullen for helpful discussions and comments on the manuscript.

-
- [1] R. Elber and M. Karplus, *J. Am. Chem. Soc.* **112**, 9161 (1990).
 [2] P. Koehl and M. Delarue, *Curr. Opin. Struct. Biol.* **6**, 222 (1996).
 [3] I. Hofacker and K. Schulten, *Proteins: Struct., Funct., Genet.* **30**, 100 (1998).
 [4] M.L. Quillin *et al.* *J. Mol. Biol.* **245**, 416 (1995).
 [5] C. Simmerling *et al.* *J. Am. Chem. Soc.* **122**, 8392 (2000).
 [6] A. Roitberg and R. Elber, *J. Chem. Phys.* **95**, 9277 (1991).
 [7] C. Simmerling, J.L. Miller, and P. Kollman, *J. Am. Chem. Soc.* **120**, 7149 (1998).
 [8] A.V. Finkelstein and B.A. Reva, *Nature (London)* **351**, 497 (1991).
 [9] A. Roitberg and R. Elber *J. Chem. Phys.* **95**, 9277 (1991).
 [10] P. Koehl and M. Delarue, *J. Mol. Biol.* **239**, 249 (1994).
 [11] C. Lee *J. Mol. Biol.* **236**, 918 (1994).
 [12] Q. Zheng and D.J. Kyle, *Proteins: Struct., Funct., Genet.* **19**, 324 (1994).
 [13] M. Vasquez *Biopolymers* **36**, 53 (1995).
 [14] Q. Zheng *et al.*, *Protein Sci.* **3**, 493 (1994).
 [15] P. Koehl and M. Delarue, *Nat. Struct. Biol.* **2**, 163 (1995).
 [16] B.A. Reva and A.V. Finkelstein, *Protein Eng.* **5**, 625 (1992).
 [17] H. Kono and J. Doi, *Proteins: Struct., Funct., Genet.* **19**, 224 (1994).
 [18] S. Kirkpatrick, C.D. Gelatt, Jr., and M.P. Vecchi, *Science* **220**, 671 (1983).
 [19] C. Tsallis and D.A. Stariolo, *Physica A* **233**, 395 (1996).
 [20] G.A. Huber and J.A. McCammon, *Phys. Rev. E* **55**, 4822 (1997).
 [21] L. Piela, J. Kostrowicki, and H.A. Scheraga, *J. Phys. Chem.* **93**, 3339 (1989).
 [22] T. Huber, A.E. Torda, and W.F. van Gunsteren, *J. Phys. Chem. A* **101**, 5926 (1997).
 [23] J. Ma, D. Hsu, and J.E. Straub, *J. Chem. Phys.* **99**, 4024 (1993).
 [24] C.D. Maranas and C.A. Floudas, *J. Chem. Phys.* **100**, 1247 (1994).
 [25] T. Huber and W.F. van Gunsteren, *J. Phys. Chem. A* **102**, 5937 (1998).
 [26] T. Huber, A.E. Torda, and W.F. van Gunsteren, *Biopolymers* **39**, 103 (1996).
 [27] J.E. Straub and M. Karplus, *J. Chem. Phys.* **94**, 6737 (1991).
 [28] A. Ulitsky and R. Elber, *J. Chem. Phys.* **98**, 3380 (1993).
 [29] W. Zheng and Q. Zheng, *J. Chem. Phys.* **106**, 1191 (1997).
 [30] C. Stultz and M. Karplus, *J. Chem. Phys.* **109**, 8809 (1998).
 [31] A. Ulitsky and R. Elber, *J. Phys. Chem.* **98**, 1034 (1994).
 [32] C. Simmerling, T. Fox, and P. Kollman, *J. Am. Chem. Soc.* **120**, 5771 (1998).
 [33] G. Verkhiver, R. Elber, and W. Nowak, *J. Chem. Phys.* **97**, 7838 (1992).
 [34] R. L. Liboff, *Kinetic Theory: Classical, Quantum, and Relativistic Descriptions* (J. Wiley, New York, 1998); I. Prigogine, *Non-equilibrium Statistical Mechanics* (Interscience, New York, 1962).
 [35] R.B. Gerber, V. Buch, and M.A. Ratner, *J. Chem. Phys.* **77**, 3022 (1982).
 [36] This was obtained via the identity $\delta(x-x')\delta(y-y') = [1/(|J|)]\delta(\alpha-\alpha')\delta(\beta-\beta')$ c.f. in CRC Standard Mathematical Tables and Formulae, 30th ed. (CRC Press, Boca Raton, FL, 1996).
 [37] H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA, 1950) pp 69-71.