JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Efficient Sampling of a Dual-Resolution Ensemble by Means of Dragging

Elias Alphonsus Jozef Franciscus Peters* and Gijsbertus de With

Laboratory of Materials and Interface Chemistry, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Supporting Information

ABSTRACT: A method to simulate a dual-resolution ensemble for molecular systems is introduced. The dual-resolution system is characterized by an atomistic Hamiltonian and coarse coordinates connected by linear springs to this atomistic system. A 'dragging' update scheme based on an idea of Neal (Neal, R. M. *Taking Bigger Metropolis Steps by Dragging Fast Variables*; Technical Report; University of Toronto: Toronto, Canada, October, 2004; http://arxiv.org/PS_cache/math/pdf/0502/0502099v1.pdf) is proposed. It is theoretically proven that the scheme correctly samples the dual ensemble. As a proof-of-principle we show that in an one-dimensional barrier crossing simulation, the relaxation speeds up by a factor 80. In an asymmetric two-dimensional barrier crossing problem, the speedup is a factor 20. The application to molecular simulations is discussed.

1. INTRODUCTION

In this paper we consider the sampling of an equilibrium distribution of a molecular system. Computational methods, such as Monte Carlo sampling or molecular dynamics, can be inefficient due to the roughness of the energy landscape. A system might get stuck for a long time in a local energy minimum. A way to reduce the problem is to consider a coarse-grained description where many degrees of freedom are replaced by a single one.

An example of coarse graining is the use of united atoms instead of a fully atomistic description.[2] More progressive coarse grainings, i.e., beyond the united atom model, are also commonly applied nowadays (see, e.g., Marrink et al., ref 3). In the coarse-grained simulation the energy landscape will be much smoother as compared to the fine-grained one, and thus barriers can be more easily crossed. Besides this, there are less degrees of freedom to be simulated. All-in-all the process of equilibration in such a model will be much faster.

Accurate coarse-grained descriptions are hard to obtain. The results of a coarse-grain simulation will deviate from the underlying, fine grain, results. The loss of fine-grain information can be problematic if one is interested in the details. A strategy to overcome these drawback is to reintroduce fine-grain information into the coarse model. This reintroduction in turn induces its own set of problems. The problems are of two kinds, namely, efficiency and accuracy. One wants to spend as little as possible computational resources on the reintroduction. Ideally the state after the fine-graining procedure, i.e., the reintroduction, should be an equilibrium state. If the state is far from equilibrium, then computation time is spent on equilibrating the fine-grained system.

Reintroduction of the fine-grained information in molecular simulations is being actively researched.[4] Here we will highlight a recently proposed class of methods closely related to our method in what it aims to achieve. The first such method was named dual-resolution replica exchange.[5] Somewhat later a more efficient method called resolution exchange (ResEx) was proposed.[6] At present, the term 'ResEx' is used for a family of similar methods.[7]

ResEx methods consider an extended ensemble. In the dual case there are two subsystems, namely, the coarse-grain system characterized by an approximate coarse-grained Hamiltonian and the fine-grain atomistic model. The extended system exists of both the coarse- and fine-grain systems combined. The equilibrium distribution of the combined system is the product of the distributions of the coarse- and fine-grain systems. This means the different levels of description are considered as statistically independent.

The general idea of exchange simulations is to exchange values for variables between the subsystems in such a way that the equilibrium distribution of the extended system remains invariant. A well-known example of such a method is parallel tempering[8] in which the different subensembles have different temperatures and the positions of the particles are exchanged between the subsystems in equilibrium at different temperatures. For this approach Monte Carlo moves are accepted using a Metropolis rule.

In the case of ResEx the fine-grain coordinates are decomposed into coarse coordinates and high-resolution ones. The Monte Carlo move consists of exchanging the coarse part of the coordinates in the fine-grain system with the coordinates in the coarse-grained system. The method has some obvious difficulties associated with it.

Coarse coordinates are straightforward to define, e.g., as center-of-mass position of groups of atoms, while the high-resolution (fine-grain) coordinates relative to a center-of-mass are difficult to parametrize. The difficulty lies in defining them in such a way that, after exchange of the coarse-grained coordinates, the reconstructed fine-grain conformation is still a likely equilibrium conformation.

For the dual-resolution method[5] improvements of the reintroduction of fine-grain information can be found in Liu and Voth[9]

Received: February 2, 2011
Published: July 28, 2011

and Liu et al.[7] When the high-resolution coordinates are not relaxed, the conformation formed by an exchange will be far from equilibrium and will have a prohibitively low acceptance probability. A strategy to increase acceptance rates in ResEx is a construction of a ladder of intermediate systems.[10] Each subsystem is only a partial coarse graining of the level below. This means that the extended system is much expanded.

Motivated by the difficulties of exchanging the high-resolution degrees of freedom in ResEx, we propose an alternative approach. We will also make use of an extended ensemble. The fine-grain system is again determined by the atom positions. The coarse-grain beads are connected by means of linear springs to the atoms. This means that the total Hamiltonian of the extended system is the atomistic Hamiltonian, containing fine-grain variables only, plus a linear spring contribution coupling the fine-grain variables to the coarse-grain ones. Different from exchange ensembles, the systems are not statistically independent. However, after integration over the coarse-grain variables, the atomistic equilibrium distribution is recovered.

Instead of an exchange mechanism we will adopt a dragging scheme as proposed by Neal.[1] We will let the coarse-grained system evolve according to an approximate coarse-grain Hamiltonian. The fine-grain system is initially connected by linear springs to the old positions of the coarse system. Next, a set of springs is also connected to the new coarse coordinates. Initially these springs are of strength zero. The springs connecting the system to the old positions will be loosened, and the springs connecting it to the new positions tightened. The springs combined with the atomistic Hamiltonian drive the dynamics. Having created a new fine-grain configuration, we will accept it or not. Using the recipe of Neal this acceptance probability can be chosen such that detailed balance is obeyed, and the correct ensemble is sampled.

The paper is organized as follows. First our dual-resolution ensemble is introduced. Next, we provide the recipe for the dragging scheme and the proof it obeys detailed balance conditions. As a proof-of-principle, we will provide simulations on a simple one- and two-dimensional system. We will finalize with a discussion and an outlook on application to molecular simulations.

2. THEORY AND METHOD

The goal is to efficiently sample a canonical ensemble associated with a fine-grain, molecular system. This density associated with the fine-grain canonical ensemble is given by

$$\rho(\Gamma) = \frac{f(\Gamma)}{Z}, \text{ where}$$
$$f(\Gamma) = \exp[-\beta H(\Gamma)] \text{ and } Z = \int f(\Gamma)\mathrm{d}\Gamma \tag{1}$$

with the conventional $\beta = 1/kT$.

We will consider two levels of description, namely, a fine grain one indicated by Γ and a coarse-grain one indicated by X. In a typical molecular simulation Γ indicates positions of all atoms and possibly their momenta. A coarse graining consists of associating a lower dimensional state X to every Γ via a function $X(\Gamma)$. If the coarse graining consists of grouping atoms into effective particles, then X indicates, for example, the center-of-mass of these groups (and possibly the center-of-mass momenta).

A coarse-grained distribution is defined as

$$\rho_{\mathrm{cg}}(X) = \frac{f_{\mathrm{cg}}(X)}{Z}, \text{ where}$$
$$= \int \delta[X(\Gamma) - X] \exp[-\beta H(\Gamma)]\mathrm{d}\Gamma$$
$$f_{\mathrm{cg}}(X) \equiv \exp[-\beta H_{\mathrm{cg}}(X)] \tag{2}$$

This gives an expression for the effective coarse-grained Hamiltonian H_{cg}. If $H_{\mathrm{cg}}(X)$ is known, all kinds of thermodynamic quantities can be computed using this coarse-grained Hamiltonian, since the Helmholtz free energy is given by

$$\exp\left[-\frac{F}{kT}\right] = Z = \int f_{\mathrm{cg}}(X)\mathrm{d}X$$
$$= \int \exp[-\beta H_{\mathrm{cg}}(X)]\mathrm{d}X \tag{3}$$

One of the methods of computing the quantities is by using coarse-grain simulations where the Hamiltonian H_{cg} provides the forces or the Monte Carlo weights.

Usually the microscopic Hamiltonian $H(\Gamma)$ is a relatively simple expression, e.g., kinetic energy plus a sum of pair potentials. The nature of $H_{\mathrm{cg}}(X)$ is much more complicated. Therefore an analytic expression for H_{cg} is usually not available. In most coarse-graining procedures that are used in practice, an approximate Hamiltonian $H_{\mathrm{approx,cg}}$ is constructed that is close to H_{cg} but easier to handle. A common approach is to take $H_{\mathrm{approx,cg}}$ to consist of a kinetic energy and a pair potential part. A popular way to construct this pair potential is the inverse Boltzmann method,[11] but other approaches are also possible.[12]

The introduction of an approximate coarse-grained Hamiltonian gives rise to deviations. Furthermore one might be interested in quantities that depend on the fine-grain information. Therefore one would like to have the fine-grain system available, next to the coarse one. For this purpose we propose to sample a dual-resolution ensemble:

$$\rho_{\mathrm{dual}}(\Gamma, X) = \frac{\sqrt{(\beta/2\pi)^d \det K_0}}{Z} f_{\mathrm{dual}}(\Gamma, X), \text{ where}$$
$$f_{\mathrm{dual}}(\Gamma, X) = \exp\left[-\beta\left(H(\Gamma) + \frac{1}{2}(X - X(\Gamma))^T \cdot K_0 \cdot (X - X(\Gamma))\right)\right] \tag{4}$$

with Z the partition sum as defined in eq 1. The coarse degrees of freedom X are connected to $X(\Gamma)$ by means of linear springs with spring constants K_0. Note that integration over X gives the fine-grain canonical distribution:

$$\rho(\Gamma) = \int \rho_{\mathrm{dual}}(\Gamma, X)\mathrm{d}X \tag{5}$$

Integration of the dual distribution over the fine-grained coordinates gives

$$\rho_{\mathrm{cg}}^{(K_0)}(X) = \int \rho_{\mathrm{dual}}(\Gamma, X)\mathrm{d}\Gamma$$
$$= \sqrt{(\beta/2\pi)^d \det K_0} \int \rho_{\mathrm{cg}}(Y) \exp\left[-\frac{1}{2}\beta(X - Y)^T \cdot K_0 \cdot (X - Y)\right]\mathrm{d}Y \tag{6}$$

The distribution is a convolution of the coarse-grained distribution ρ_{cg}, given by eq 2, with a Gaussian with variance (matrix) $(\beta K_0)^{-1}$. An effective Hamiltonian can now be defined by

$$\exp[-\beta H_{cg}^{(K_0)}(X)] = \sqrt{(\beta/2\pi)^d \det K_0} \int \exp\left[-\beta\left(H_{cg}(Y) + \frac{1}{2}(X-Y)^T \cdot K_0 \cdot (X-Y)\right)\right] dY \quad (7)$$

For infinitely stiff springs, i.e., $K_0 \to \infty$, $\rho_{cg}^{(\infty)} = \rho_{cg}$ and $H_{cg}^{(\infty)} = H_{cg}$.

The procedure we propose to sample the dual-resolution ensemble is to use an approximate model to generate new coarse-grain variables X^* from the old ones X. Next, the fine-grain variables are dragged from an initial state Γ to a final state Γ^*. This is established by connecting the fine-grain state by means of springs to both X and X^* and changing the spring constants gradually. The configuration is dragged from X to X^* by loosening the springs connecting the fine-grain coordinates to X and tightening those connected to X^*. Finally the new dual-resolution state (Γ^*, X^*) is accepted or not. The acceptance rule should be such that the density described by eq 4 is sampled. The idea of *dragging* the fine-grained structure is borrowed from Neal.[1] The new elements in our approach are the addition of the extra variables X and the linear springs. A cartoon of this procedure is depicted in Figure 1.

In our computational method first X^* is generated from the previous position X. The new X^* is sampled by means of the conditional probability distribution $w(X^*|X)$. The equilibrium X needs to sample the Boltzmann distribution corresponding to the approximate coarse-grain Hamiltonian $H_{approx,cg}$:

$$\rho_{approx,cg}(X) = \frac{f_{approx,cg}(X)}{Z_{approx,cg}}, \text{ where}$$
$$f_{approx,cg}(X) = \exp[-\beta H_{approx,cg}(X)] \quad (8)$$

Therefore the detailed balance condition for $w(X^*|X)$ becomes

$$w(X^*|X)\rho_{approx,cg}(X) = w(X|X^*)\rho_{approx,cg}(X^*) \quad (9)$$

We will assume that X is generated using some Markovian simulation method [a Monte Carlo (MC) method, molecular dynamics (MD), or Brownian dyamics (BD)] that obeys this relation.

Having generated the final X^* we can start dragging the fine-grain model from X to X^*. During the dragging, the microscopic system evolves according to the Hamiltonian:

$$H_{drag}(\Gamma,\lambda) = H(\Gamma) + \frac{1}{2}K(\lambda) : (X(\Gamma) - X)^2 + \frac{1}{2}K^*(\lambda) : (X(\Gamma) - X^*)^2 \quad (10)$$

where the parameter λ is used to parametrize the magnitude of the spring constants. It is changed from 0 to 1 during the dragging procedure. We will take

$$K(0) = K_0, \quad K(1) = 0, \quad \text{and} \quad K^*(\lambda) = K(1-\lambda) \quad (11)$$

This means that for $\lambda = 0$, the fine-grain state is coupled by means of linear springs to the initial coarse-grain state X, while for $\lambda = 1$ the fine-grain state is coupled to the final coarse-grain state X^* and not connected to X anymore. The relation between K^* and K is such that the scheme can easily made to obey detailed balance.

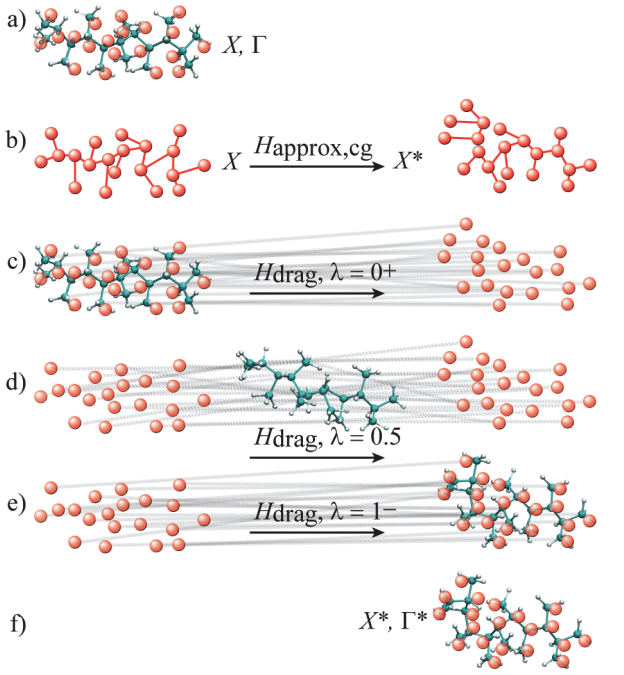

Figure 1. A cartoon of the dragging scheme. The fine-grain state Γ is depicted as a molecule. The coarse-grain state X is depicted as the larger (red) beads. These beads are connected by means of linear springs to the center-of-masses of groups of atoms (in this case carbon with its hydrogens). (a) The initial dual-resolution state consists of the atomistic model and the coarse beads connected to it. (b) The first step in the updating scheme is to move X to the new state X^* by using some approximate Hamiltonian. (c) Next the atomistic model is connected to both the old state X and the new state X^*. Each center-of-mass is connected to one bead in the old coarse-grain state X and one bead in the new coarse-grain state X^*. The magnitude of the spring constants is parametrized by a variable λ. Initially, for λ near to zero, the springs connecting the molecule to X are much stronger than the springs connected it to X^*. (d) Gradually λ is increased from 0 to 1. Simultaneously the molecule performs its natural dynamics. It is influenced by intra- and intermolecular forces as well as by the spring forces. Because the springs connected to X loosen and those connected to X^* become stiffer, the molecule will move toward X^*. (e) For λ near to 1 the molecule will be close to X^*. (f) The final dual-state is characterized by Γ^* and X^* and should be accepted or rejected according to the appropriate Metropolis rule.

At intermediate states where $\lambda \neq 0,1$, we define

$$\rho(\Gamma|X^*,X,\lambda) = c(X^*,X,\lambda) f(\Gamma,X^*,X,\lambda) \text{ with}$$
$$f(\Gamma,X^*,X,\lambda) = \exp[-\beta H_{drag}(\Gamma,\lambda)] \quad (12)$$

where $c(X^*,X,\lambda)$ is the normalization constant. This is the equilibrium distribution for the microscopic state Γ, for the case that the system would evolve according to the Hamiltonian with the value of λ fixed. In the scheme, however, the value of λ is changed from 0 to 1. Depending on the rate that λ changes, the distribution of Γ is close to $\rho(\Gamma|X^*,X,\lambda)$ or not. The fact that during the dragging the intermediate states are not in local equilibrium does not influence the validity of the scheme as will be discussed below.

The variable λ is changed in discrete steps $i = 1, ..., n$ from 0 to 1,

$$\lambda_i = i/(n + 1) \tag{13}$$

The microstate Γ_i is generated from the previous one at $i - 1$ by the transition probability $T(\Gamma_i|\Gamma_{i-1},X^*,X,\lambda_i)$. Here $\Gamma_0 = \Gamma$, and $\Gamma^* = \Gamma_n$. For each λ we take care that the conditional transition probability $T(\Gamma'|,X^*,X,\lambda)$, to go from state Γ to state Γ', obeys the detailed balance condition:

$$\begin{aligned} &T(\Gamma'|\Gamma, X^*, X, \lambda)\rho(\Gamma|X^*, X, \lambda) \\ &\quad = T(\Gamma|\Gamma', X^*, X, \lambda)\rho(\Gamma'|X^*, X, \lambda) \end{aligned} \tag{14}$$

This means that the transition probabilities can be generated using conventional simulation techniques, such as MC, MD, or BD using the Hamiltonian eq 10 with the appropriate value of λ for each step.

Finally, the path from initial state $(\Gamma_0 = \Gamma,X)$ to final state $(\Gamma^* = \Gamma_n,X^*)$ is accepted with a probability: $\mathrm{acc}(\Gamma_n, ..., \Gamma_1,X^*|\Gamma_0, X)$. The total probability density of this process becomes

$$\begin{aligned} &\rho(\Gamma_n, ..., \Gamma_0, X^*, X) \\ &\quad = \mathrm{acc}(\Gamma_n, \ldots, \Gamma_1, X^*|\Gamma_0, X) \\ &\quad \times w(X^*|X) \prod_{i=1}^{n} T(\Gamma_i|\Gamma_{i-1}, X^*, X, \lambda_i)\rho_{\mathrm{dual}}(\Gamma_0, X) \end{aligned} \tag{15}$$

By requiring that a detailed balance condition for the acceptance of the path is obeyed, we find that

$$\begin{aligned} &\frac{\mathrm{acc}(\Gamma_n, ..., \Gamma_1, X^*|\Gamma_0, X)}{\mathrm{acc}(\Gamma_0, ..., \Gamma_{n-1}, X|\Gamma_n, X^*)} \\ &\quad = \frac{w(X|X^*)}{w(X^*|X)} \prod_{i=1}^{n} \frac{T(\Gamma_{i-1}|\Gamma_i, X^*, X, \lambda_i)}{T(\Gamma_i|\Gamma_{i-1}, X^*, X, \lambda_i)} \frac{\rho_{\mathrm{dual}}(\Gamma_n, X^*)}{\rho_{\mathrm{dual}}(\Gamma_0, X)} \end{aligned} \tag{16}$$

Inserting the detailed balance conditions for the individual transition probabilities, eqs 14 and 9, results in

$$\begin{aligned} &\frac{\mathrm{acc}(\Gamma_n, ..., \Gamma_1, X^*|\Gamma_0, X)}{\mathrm{acc}(\Gamma_0, ..., \Gamma_{n-1}, X|\Gamma_n, X^*)} \\ &\quad = \frac{\rho_{\mathrm{approx,\,cg}}(X)}{\rho_{\mathrm{approx,\,cg}}(X^*)} \prod_{i=1}^{n} \frac{\rho(\Gamma_{i-1}|X^*, X, \lambda_i)}{\rho(\Gamma_i|X^*, X, \lambda_i)} \frac{\rho_{\mathrm{dual}}(\Gamma_n, X^*)}{\rho_{\mathrm{dual}}(\Gamma_0, X)} \\ &\quad = \frac{f_{\mathrm{approx,\,cg}}(X)}{f_{\mathrm{approx,\,cg}}(X^*)} \prod_{i=0}^{n} \frac{f(\Gamma_i, X^*, X, \lambda_{i+1})}{f(\Gamma_i, X^*, X, \lambda_i)} \\ &\quad = \exp\Big[\beta\{H_{\mathrm{approx,\,cg}}(X^*) - H_{\mathrm{approx,\,cg}}(X) \\ &\quad - \frac{1}{2}\sum_{i=0}^{n}((K(\lambda_{i+1}) - K(\lambda_i)) : (X(\Gamma_i) - X)^2 \\ &\quad + (K^*(\lambda_{i+1}) - K^*(\lambda_i)) : (X(\Gamma_i) - X^*)^2)\}\Big] \end{aligned} \tag{17}$$

This ratio of acceptance probabilities is used in a Metropolis rule by choosing the acceptance probability of the generated path as

$$\begin{aligned} &\mathrm{acc}(\Gamma_n, ..., \Gamma_1, X^*|\Gamma_0, X) \\ &\quad \equiv \min\left(1, \frac{\mathrm{acc}(\Gamma_n, ..., \Gamma_1, X^*|\Gamma_0, X)}{\mathrm{acc}(\Gamma_0, ..., \Gamma_{n-1}, X|\Gamma_n, X^*)}\right) \end{aligned} \tag{18}$$

This might seem a circular definition, but it is not. Clearly if one computes the ratio of these acceptance probabilities for the two paths, it will equal the ratio in the right-hand side, so the definition is consistent. It is, however, a further specification of the acceptance probability which is now, by construction, guaranteed to lie between 0 and 1. For the ratio appearing at the right-hand side, eq 17 needs to be substituted thus giving a closed definition for computing the acceptance probability.

The final result of the scheme is the generation of a path with probability density $\rho(\Gamma_n, ..., \Gamma_0,X^*,X)$, that gives rise the desired dual-resolution distribution for Γ^* and X^*, eq 4:

$$\rho_{\mathrm{dual}}(\Gamma^*, X^*) = \int \rho(\Gamma^*, ..., \Gamma_0, X^*, X)\mathrm{d}\Gamma_{n-1}, ..., \mathrm{d}\Gamma_0\mathrm{d}X \tag{19}$$

A derivation of the main result eq 16 is given in the Appendix A. The derivation is based on the work of Neal.[1]

It is probably worthwhile to point out some of the subtleties that might be overlooked otherwise. In the scheme, the index, i, on λ_i runs from 0 to $n + 1$ (including end points), while the index on Γ_i runs from 0 to n. If this were not the case, then the ratio of transition probabilities in eq 16 and the ratio of probability distributions in eq 17 normalization constants would not cancel. This means that for the transition Γ to Γ_1, the springs connecting Γ to X^* are already turned on by a small amount. Likewise for the last step Γ_{n-1} to Γ^*, the springs connecting Γ to X are still weakly turned on. These details are important for implementing the scheme, especially if n is not large.

In a dynamic simulation the time interval corresponding to a change of λ does not have to correspond to an integration time step of the equation of motion. One is allowed to divide a constant λ_i interval into more subintervals. Such a subdivision might be needed because the time step for integrating the equations of motion needs to be sufficiently small. When subdividing the constant λ intervals, some care has to be taken so that eq 35 remains valid.

Although eq 17 is expressed in terms of equilibrium distributions for Γ, this equilibrium does not have to be reached. The dragging is allowed to be performed out of equilibrium. The intermediate distributions in the expressions only arise because of eq 14. This gives a large flexibility to designing possible dragging schemes.

Equation 16 is a more generally valid expression than eq 17 since it can be used even if eq 14 is not valid. In many dynamical simulation methods, such as MD and BD, eq 14 is not exactly obeyed due to time discretization errors. In the BD simulation method the transition probabilities can often be exactly determined. In this case one can use eq 16 and prevent time discretization errors.

3. ANALYSIS OF METHOD AND SIMULATION RESULTS

In the previous section we gave a theoretical derivation of the dragging scheme. It remains to be proven that the scheme is useful in practice. We will consider some simple systems to provide a prove of principle. Besides this we want to develop an intuition for choosing the parameters of the scheme.

When the dragging has been performed, the newly generated state is accepted or rejected. For the scheme to be useful in practice, the acceptance probability should be reasonable. In Appendix B we consider the case of very slow dragging. If the number of subdivisions n is large enough and if also the approximate coarse-grain Hamiltonian $(H_{\mathrm{approx,cg}})$ is close to the real smoothed one $(H_{\mathrm{cg}}^{(K_0)})$, the acceptance ratio approaches 1.

The situation described in Appendix B is the optimal case. If the approximate Hamiltonian is not a good approximation or if the dragging is not slow, then the acceptance probabilities will

most likely drop. This is numerically investigated in the next two subsections. In both subsections a barrier crossing problem is considered. We think this is the archetypal situation where the dragging scheme can be useful. The barrier is so high that crossing it becomes a rare event when using conventional methods to simulate the dynamics.

The approximate Hamiltonian in the dragging scheme will be chosen such that the coarse-grain variable can relatively easily cross the barrier. By making the springs that connect the fine-grain variables to the coarse-grain ones stiff enough, the fine-grain variables can be pulled over the barrier. The hope is that parameters of the dragging scheme can be chosen in such a way that the acceptance probability is reasonable. If this is the case, then it means that barriers are crossed more easily. Since the barrier crossing is the rate-determining step, the system will equilibrate faster.

In both problems BD is used for generating the positions.[13,14] In BD only the particle positions are considered, not the momenta. The displacements of a particle have a fluctuating (stochastic) part that is characterized by the diffusion coefficient and a deterministic part that is driven by a potential U. Only the configurational part of the canonical distribution is relevant for BD. Therefore the potential U will play the same role as the Hamiltonian used in the general derivation of the previous section. This would also be the case for MC simulation. We choose BD, and not MC, because there is a natural time scale which makes it more straightforward to quantify the speedup of the dragging scheme.

3.1. The Double-Well Problem. The simplest barrier crossing problem is the one-dimensional double-well problem. We use the potential:

$$\beta U_h(x) = x^2(h\cdot(2x^2-3)x^2 - 2(1-x^2)^2) \qquad (20)$$

Here h is the barrier height expressed in units kT. The rationale for the functional form chosen is that the barrier height is exactly h, the wells are located at $x = -1$, a $x = 1$ irrespective of the barrier height and the curvature at the barrier is independent of h. In Figure 2 the potential is drawn for $h = 10$, which is the case studied. In the lower graph of the same figure we show the corresponding canonical equilibrium distribution. Also a simulation result obtained using the dragging scheme is shown.

In the one-dimensional case there is no real coarse-graining taking place. We do, however, introduce an approximate coarse-grain potential that has the same functional form as the fine-grain one, eq 20, but with a different barrier height. Values of $h = 1-4$ are considered. The reason is that now the coarse-grain position X will quite frequently cross its barrier. The starting and final coarse-grain positions, X and X^*, are connected to the fine-grain position, x, by means of two linear springs. The hope is that x will be dragged over the barrier and that the new configuration is accepted.

For the presented results both the units of length and time are taken to be 1. The energy scale is kT. The BD simulations of both x and X are performed with a diffusion coefficient of 1. The new position X^* is taken at a time $\Delta t = 5$ from the starting position X. The λ's are updated in $n = 100$ discrete steps during this time interval (see eq 13). The spring constants are varied linearly as $K(\lambda) = K_0\cdot(1-\lambda)$. The time step used to integrate the equations within the BD algorithm is a fraction of 0.05 such that discretization errors are negligible.

3.1.1. The Stuck Problem. In the upper graph of Figure 3, a typical time series of the dragging scheme is shown. A point is plotted once so many time steps. Initially the coarse-grain

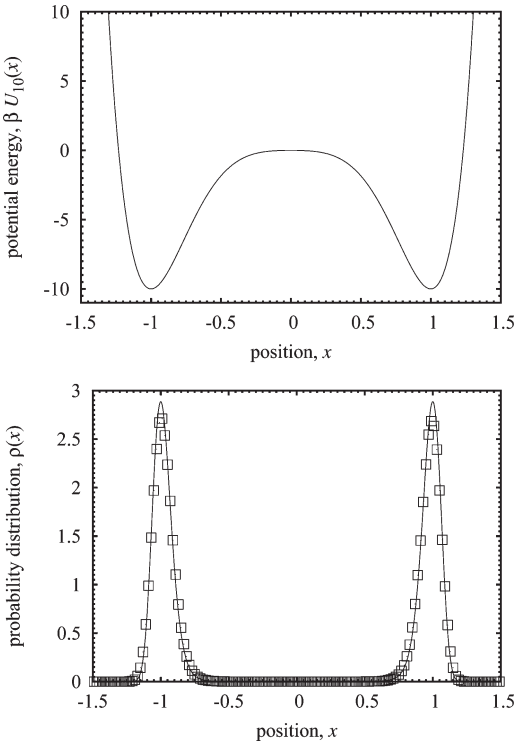

Figure 2. The upper graph shows the used double-well potential with a barrier height of $10kT$. The lower graph shows the corresponding canonical distribution. The symbols are the histogram found by means of the dragging scheme.

variable frequently hops from the $X = -1$ to the $X = 1$ position. The fine-grain variable x follows along. Quite often in the time series the positions X and x get stuck for a long time at a certain position. Such a position is typically far from the potential minimum. The simulation gets stuck because the acceptance ratio, eq 17, drops to nearly zero. The cause is the presence of the factor $\exp[\beta(H_{\text{approx,cg}}(X^*)]$ in the acceptance ratio. This factor is included to cancel the influence of the approximate coarse-grained dynamics on the final statistics.

In the Supporting Information a document is included in which this problem is analyzed. It turns out that the factor:

$$\exp[\beta(H_{\text{approx, cg}}(X^*) - H_{\text{cg}}(X^*))] \qquad (21)$$

determines the observed stuck behavior. If for large X^* this factor becomes very large, then this means that the approximate dynamics under samples this state compared to the real dynamics. To compensate for the fact that such a state is very rarely visited, the residence time, once the state is visited, should become very large. This means the simulation is stuck for a very long time. This gives rise to an ergodicity problem. The simulation time should be long compared to the time states are stuck. This time can be prohibitively long.

There is a simple remedy for this problem. An alternative way to sample the dual-resolution ensemble, eq 4, is to generate fine-grain states Γ (i.e., x for the system under consideration) and then connect the coarse-grain states X to it by means of drawing a Gaussian distributed displacement with a variance $(\beta K_0)^{-1}$. The sampling of Γ can straightforwardly be done by letting this state evolve according to its dynamics. If we apply this 'move' subsequent to a 'dragging move', then the coarse-grain position

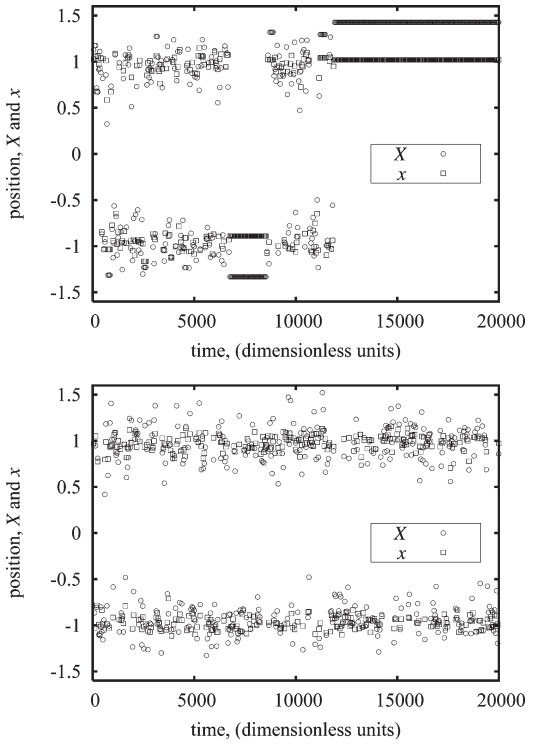

Figure 3. Time series of fine-grain positions x and coarse-grain positions X. The upper graph shows a time series generated by only using the dragging scheme. Configurations have the tendency to get stuck. In the lower graph, the dragging scheme and ordinary dynamics (with reconstruction of the coarse variables) are alternated. In this case configurations do not get stuck. The histogram in Figure 2 corresponds with this time series.

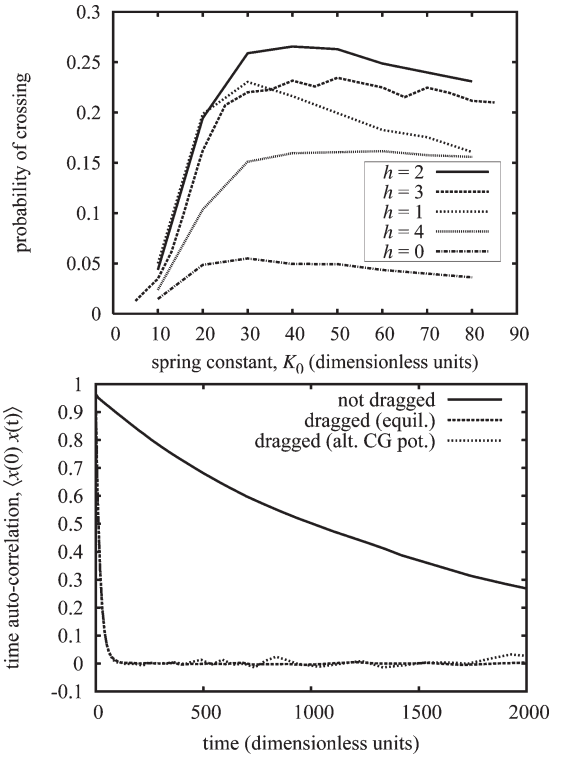

Figure 4. In the upper graph, crossing probabilities as a function of the spring constant K_0 are plotted. The crossing probability is the probability that the final positions, x, are at the other side of the $10kT$ barrier. The potentials used to generate coarse-grain position, i.e., eq 20, are labeled by their potential barrier height, h (in units kT). The lower graph shows the time autocorrelation of position, x. For the normal Brownian motion the decay is slow because a barrier of height $10kT$ needs to be crossed. For a dragged simulation, with well chosen parameters, the decay is much faster. Both the optimized potentials of functional forms eqs 20 and 22 give the same fast decay.

X will relax. Therefore the next application of the dragging scheme will not get stuck. The second graph in Figure 3 displays a time series generated by this method showing that configurations are no longer stuck, but both sides of the barrier are evenly sampled. The histogram shown in Figure 2 corresponds to this simulation.

In the example depicted in Figure 3 the approximate coarse-grained dynamics was generated with the potential eq 20 (with a low barrier). For large X, this potential diverges as $U_{\text{approx,cg}}(X) \propto X^6$. For the dual-resolution ensemble, eq 4, we expect that Γ is near one of the wells and that X is fluctuating in a Gaussian cloud around this position. Therefore we expect $U_{\text{cg}}(X) \propto X^2$. Clearly the factor given by eq 21 is very divergent in this case. This explains why the pure dragging simulation gets stuck. According to the reasoning this stuck problem can also be resolved by choosing $U_{\text{approx,cg}}(X)$ in such a way that eq 21 is well behaved. To test this, we also performed simulations with an approximate coarse-grained potential of the form:

$$\beta U_{\text{approx, cg}}(X) = -\ln\left(\exp\left[-\frac{\tilde{K}}{2}(X+1)^2\right] + \exp\left[-\frac{\tilde{K}}{2}(X-1)^2\right] + C\,\text{erfc}\left[\sqrt{\frac{\tilde{K}}{2}}(-X-1)\right]\text{erfc}\left[\sqrt{\frac{\tilde{K}}{2}}(X-1)\right]\right) \quad (22)$$

Here $\tilde{K}$ determines the curvature at the potential well. The error functions part can be used to lower the barrier by tuning the constant C. This potential increases as X^2 for large X. We indeed found that by using this potential for the approximate coarse-grained dynamics, the stuck problem does not occur (see below).

3.1.2. Tuning Parameters. One of the main parameters to be chosen when implementing the dragging scheme is the spring constant K_0. In the upper graph of Figure 4 the average probability of crossings is plotted as function of K_0. This is the probability that x and x^* lay on different sides of the barrier subsequent to a dragging move. The reason for considering this probability is that the barrier crossing is the rate-determining step for equilibration.

The acceptance probabilities will be such that the dual-resolution distribution, eq 4, is sampled. This means that if K_0 is large enough, then x lays almost always on the same side of the barrier as X. The probability that x crosses the barrier equals the probability that X crosses the barrier times the probability that x follows. This last probability is the acceptance probability, eq 18. In the case that X and the new position X^* are statistically independent, the probability that X crosses the barrier equals 0.5. Therefore 0.5 is an upper bound for the probability of crossing.

The acceptance probability can be of a reasonable magnitude only if K_0 is so large that x can be dragged over the barrier. This means that, approximately, $1/2 \cdot K_0 \cdot 1 > 10$, so

$$K_0 > 20 \quad (23)$$

We also expect that the dragging scheme works best if the potential that is used to sample X (i.e., eq 20 with, e.g., $h = 2$) resembles the smoothed coarse-grained potential of X, eq 6. In Appendix B we show that in the limit of slow dragging, the acceptance ratio approaches zero if the approximate Hamiltonian is close to the smoothed one. For the minimum of the local well, eq 20, we can approximate the region around a minimum by a harmonic potential. A spring constant $K_{well}(h) = 8\cdot(3h - 2)$ corresponds to this harmonic potential. The smoothed coarse-grain potential, $U_{cg}^{(K_0)}$, can be locally approximated by a harmonic potential with spring constant $K_{cg,well}^{(K_0)}$. This smoothed potential is obtained by coupling a spring to position x. The strength of the effective linear springs then follows as

$$\frac{1}{K_{cg,\,well}^{(K_0)}} = \frac{1}{K_{well}(10)} + \frac{1}{K_0} \tag{24}$$

By equating this spring constant to the one associated with the potential used to generate positions X, i.e., $K_{well}(h)$, an estimate for the optimal value of K_0 can be computed. For varying values of h we find: $h = 1 \rightarrow K_0 = 8.3$, $h = 2 \rightarrow K_0 = 37.3$, $h = 3 \rightarrow K_0 = 74.7$, and $h = 4 \rightarrow K_0 = 124.4$.

The probability for X crossing the barrier is larger if its barrier is lower. For low values of h, however, the optimal value of K_0 computed with eq 24 is too small to drag x over the barrier (see eq 23). Therefore the optimal value of h is expected to be a value that is just large enough such that the optimal K_0 can drag the fine-grain variable over the barrier.

In the lower graph of Figure 4, we find an optimal value of $K_0 \approx 40$ and a barrier of height of $2kT$ for the approximate dynamics of X using the potential eq 20. The probability for crossing is about $p_{cross} = 0.26$. The dragging of x from X to X^* is performed in a time $\Delta t = 50$. Slow dragging improves the acceptance probability. For very slow dragging, we find that the crossing probability increases to 0.31. For quicker dragging, $\Delta t = 10$, we have $p_{cross} = 0.17$, and $\Delta t = 5$ gives $p_{cross} = 0.11$.

The speedup in the computation due to the dragging scheme can be determined by comparing the CPU time needed for a conventional simulation and for a dragging simulation to approach equilibrium. In 'real' molecular simulations, the overhead due to the coarse-grain dynamics will be negligible. Also the extra spring forces will contribute little. Therefore the CPU time needed to equilibrate a system will be proportional to the simulated physical time of the fine-grain dynamics.

For determining the speedup in the equilibration due to dragging, we compare the time autocorrelation of the position x. The time plotted on the x-axis of the lower graph in Figure 4 is the physical time of the fine-grain dynamics. Time runs on irrespective of whether moves are accepted or not. The time autocorrelation of the normal BD is governed by the barrier crossing. Because the barrier height is $10kT$, this is a highly activated process. For the normal dynamics, we find a decay rate of 6.4×10^{-4}, while for the dragged simulation, with dragging time $\Delta t = 5$, the decay rate is 5.2×10^{-2}, so the speedup factor is 81. Note that the decay rate is expected to be proportional to $p_{cross}/\Delta t$. Since p_{cross} decreases with decreasing Δt, there is an optimal value for Δt. We found this value to be around $\Delta t = 5$.

When using eq 22 we see the same fast decaying behavior when we use $K_0 = 40$, $\tilde{K} = 20$, and $C = 0.04$ (potential barrier of about $1.9kT$). This shows that the behavior is quite robust. It does not seem to depend much on the details of the coarse-grained potential (except for the occurrence of the stuck problem). When independently varying K_0 between 25 and 60, $\tilde{K}$ between 10 and 50, and C between 0.01 and 0.25, we found essentially the same fast decaying behavior as depicted in Figure 4. The spring strength K_0 could be increased up to 120 if we simultaneously adapted $\tilde{K}$ such that eq 24 was obeyed. For parameters that deviated more, a noticeable deviation from the fastest possible decay was observed but still the dynamics was much faster than the nondragged case.

3.2. A Two-dimensional Case. In this section we will provide a proof-of-principle for a simple two-dimensional case. We will use the example also used in refs of Lyman and Zuckerman[10] and Liu and Voth[9] for proofs-of-principle. Compared to the one-dimensional example above, there are some additional features: In this case there will be a genuine coarse graining, namely, from two- to one-dimensional. This two-dimensional example has an asymmetric fine-grain potential and a symmetric approximate coarse-grain potential $U_{approx,cg}$.

The system considered is the canonical equilibrium corresponding to the potential energy:

$$\beta U(x,y) = 10(x^2 - 1)^2 + \frac{y^2}{1 + 250(\tanh(10x) + 1)} \tag{25}$$

In this potential there are two local minima, namely at $(x,y) = (-1,0)$ and $(x,y) = (1,0)$. At $(x,y) = (0,0)$ there is a saddle point. The height of the potential barrier to go from one minimum to the other one is $10kT$. Due to the presence of the 'tanh', the basin centered around $(1,0)$ is much more extended in the y-direction than the one centered around $(-1,0)$. The combination $\Gamma = (x,y)$ signifies the fine-grain state. The coarse graining is simply defined by $X(\Gamma) = x$. The coarse-grained potential, using eq 2, becomes

$$\beta U_{cg}(X) = 10(X^2 - 1)^2 - \frac{1}{2}\ln(1 + 250(\tanh(10X) + 1)) - \frac{1}{2}\ln \pi \tag{26}$$

To generate the dynamics for X, we use the same potential as in ref of Liu and Voth,[9] namely,

$$\beta U_{approx,\,cg}(X) = (X^2 - 1)^2 \tag{27}$$

This approximate potential does have a barrier that is much lower than the (smoothed) coarse-grained potential barrier and is therefore much easier crossed. Graphs of the canonical probability densities corresponding to the smoothed version of U_{cg} (with $K_0 = 10kT$) and eq 27 are shown in Figure 6.

As a method of simulation, BD is used with diffusion coefficients $D = 1$ for both the coarse- and fine-grain dynamics. The time interval from X to X^* is $\Delta t = 2$. The same time interval is used for the fine-grain dynamics.

For the spring constants we use

$$\beta K(\lambda) = 10(1 - \lambda) + 60\lambda(1 - \lambda) \tag{28}$$

The equilibrium distribution of X is governed by $K_0 = K(0) = 10kT$. This value of K_0 broadens the canonical distribution corresponding to the 'real' coarse-grained potential U_{cg}, eq 26, such that its peaks have widths comparable to the distribution associated with eq 27. Application of eq 24 gives a value of $K_0 = 8.9kT$. The value of $K_0 = 10kT$ is too small to drag the fine-grain position over the barrier. By adding the quadratic part we find that halfway $K(1/2) = 20kT$, which is strong enough. By means of

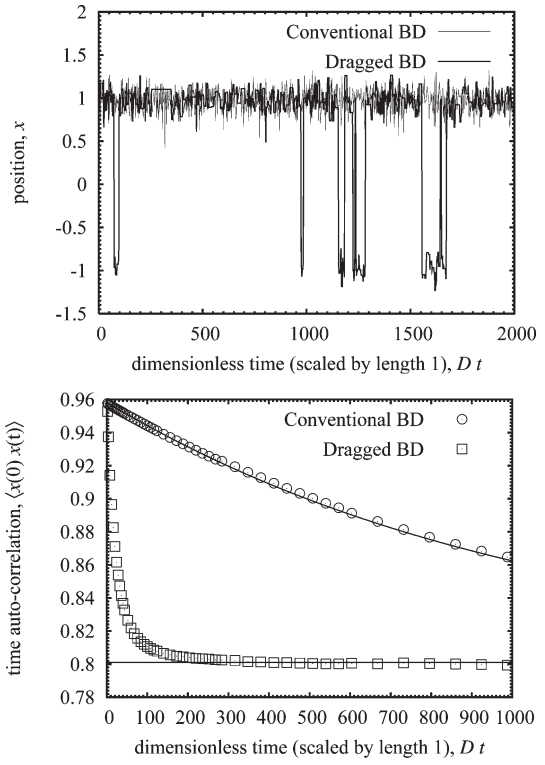

Figure 5. The upper graph shows a time series of the x-coordinate of a particle in the two-dimensional two-well problem. For conventional BD (thin line), the particle fluctuates around $x = 1$. For the dragging scheme (bold line), acceptance is lower, and the graph is sometimes flat. The dragging BD enhances barrier crossings, i.e., jumps to $x = -1$. The lower graph gives the autotime correlation for the time series. The simulation results of the dragging scheme relax much quicker to the long time limit value of 0.801.

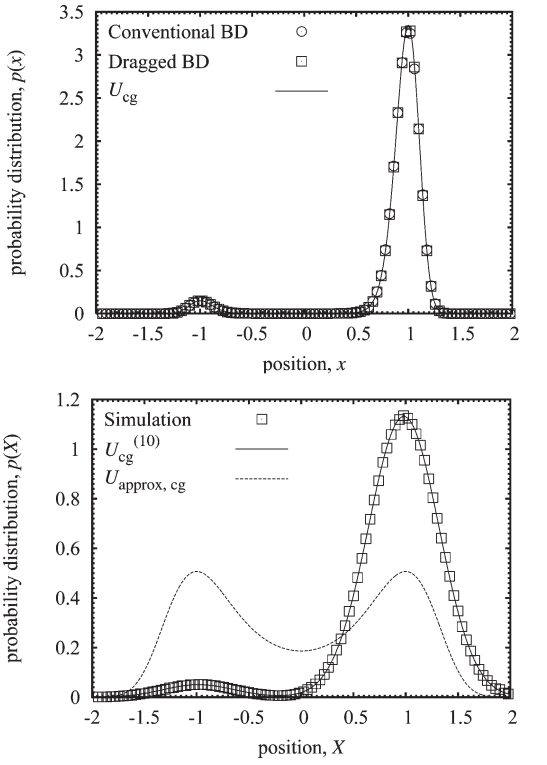

Figure 6. The upper graph shows the probability distribution of x as found in simulations. The dragging scheme samples the canonical distribution corresponding to U_{cg}, eq 26, to high accuracy. The lower graph shows the distribution for X. It samples the distribution corresponding to $U_{cg}^{(K_0)}$ with $K_0 = 10$ as it should. This distribution significantly deviates from the canonical distribution, eq 27, used to generate the moves.

trial-and-error, we determined that the choice eq 28 is close to optimal.

The upper graph in Figure 5 shows a sample time-series of positions $x(t)$. The thin line indicates simulation by conventional BD for the potential $U(x,y)$ given by eq 25. The initial position is $x = 1$ (and $y = 0$). Within the time interval simulated, the system does not cross the potential barrier. Only fluctuations around $x = 1$ are seen. For the dragging scheme, the time series is indicated by the bold line. This bold line does cross over to the local minimum around $x = -1$ several times.

The lower graph shows the time autocorrelation for $x(t)$. For $t = 0$, the equilibrium value is $\langle x^2 \rangle = 0.972$ according to eq 26. A small time later, local equilibrium has been established around $x = -1$ and 1. The equilibrium probability for $x < 0$ is $P_L = 0.0428$, while for $x > 0$ it is $P_R = 0.957$. For the average position inside the peak we find $\langle x \rangle_L = -0.979$ and for $\langle x \rangle_R = 0.979$. After local equilibration around the potential minima this gives

$$\langle x(0)x(0^+) \rangle \approx P_L \langle x \rangle_L^2 + P_R \langle x \rangle_R^2 = 0.958 \quad (29)$$

This is indeed the initial value seen in the time-correlation graph of Figure 5. For very long times, the barrier is crossed many times, so

$$\langle x(0)x(\infty) \rangle = \langle x \rangle^2 = (P_L \langle x \rangle_L + P_R \langle x \rangle_R)^2 = 0.801 \quad (30)$$

This value is indicated by means of a line in the graph.

For conventional BD, the decay of the time correlation from 0.958 to 0.801 is estimated to be a single exponential with a rate of 9.405×10^{-4}. This value is not fitted but computed by numerical determination of the lowest nonzero eigenvalue of the diffusion problem corresponding to the two-dimensional Brownian motion. The time-correlation graph shows that with the dragging scheme, decorrelation is about a factor 20 faster.

In Figure 6 the probability distributions for x and X, as obtained by the simulations, are given. The upper figure shows the histogram for x. The theoretical distribution is the canonical distribution corresponding to eq 26. Correspondence between simulation and theory is nearly perfect. The positions X are generated using the potential eq 27. The canonical probability distribution corresponding to this potential is also shown in the lower graph. The coarse variable X is, however, expected to be distributed according to eq 26 smoothed with $K_0 = 10kT$ as given by eq 6. It is clear that the difference between the 'generating' distribution and the sampled one is significant. The simulation results show that $\rho_{cg}^{(10)}$ is sampled accurately.

4. DISCUSSION AND CONCLUSIONS

We introduced a dual-resolution ensemble and implemented the Neal dragging idea for this ensemble. The dragging scheme was shown to significantly increase the efficiency of sampling the equilibrium distribution in a one- and two-dimensional potential barrier crossing problem. In the one-dimensional a large speedup factor of 81 was found, while in the two-dimensional case the speedup is a factor 20. When the scheme is generalized to

molecular simulations, this might give a significant speedup for these simulations also.

It was shown theoretically that the dragging scheme is sound, i.e., that the canonical ensemble remains invariant. This was achieved by applying the Metropolis rule such that the correct detailed balance condition for motion along a path is obeyed. In the two-dimensional example a symmetric Hamiltonian is used to generate the dynamics of the coarse-grain state. The 'real' coarse-grained Hamiltonian found by integrating out the fine-grain variables from the dual-resolution ensemble, however, is asymmetric. The corresponding asymmetric ensemble for X is well sampled by the scheme. This provides a practical example of the validity of the scheme. As a remark on the side, note that the smaller speedup factor of 20, as compared to 81, is partly due to the large mismatch between coarse-grain Hamiltonian used to generate the X and its true Hamiltonian. The scheme needs to correct for this mismatch by rejecting some of the moves. Because the potential generating the dynamics of X was given, the speedup could not be fully optimized. We used the less than optimal coarse-grain potential because it was also used proof-of-principle of ResEx methods.[9,10]

By means of analysis of the simulations of the simple systems we identified and also resolved a weakness of the dragging scheme. The problem is that the simulations get stuck because acceptance rates drop to zero. An analysis of this problem is available in the Supporting Information. It occurs if the approximate coarse-grain dynamics very rarely samples a region that should be sampled much more frequently, according to the real coarse-grain dynamics. One way to resolve the problem is to make sure this situation does not arise by choosing a suitable approximate coarse-grain Hamiltonian. It is not sure if this is always possible in practice.

A second way to resolve this stuck problem is by equilibrating the fine-grain system and regenerating the Gaussian springs. In this case the implementation of the dragging scheme alternates between two *moves*. In the dragging move the coarse-grain variables X move first and by means of dragging guide to fine-grain state Γ to a new favorable state. The second move is an equilibration move where the fine-grain system evolves according its natural dynamics. The coarse-grain variables are reconnected to the fine-grain state at the end. Both moves sample the dual-resolution ensemble correctly. By combining them, the system does not get stuck.

An envisioned application is to speedup molecular simulations. The idea is to use mesoscopic scale (coarse-grain) simulations to guide atomistic systems to sample phase space more efficiently. The mesoscopic system has the essential features of the molecular system. Configurations are, on a qualitative level, correct. Interactions are softer, and therefore, the barriers that determine the rate of equilibration are more easily crossed. A drawback of the mesoscopic level simulation is its lower accuracy due to the loss of atomistic detail. By dragging a atomistic model toward the mesoscopic configuration, it is force to cross the (higher) potential barriers. This means a speedup compared to atomistic simulation without dragging. The sampling of atomistic level means accuracy is high.

In the proof-of-principles and also in Appendix B, we found that the potentials used to generate the coarse-grain dynamics, $U_{\text{approx,cg}}$, should be close to $U_{\text{cg}}^{(K_0)}$. Most mesoscopic simulation techniques use approximate Hamiltonians that are aimed to be close to the real coarse-grained Hamiltonian defined by eq 2 corresponding to $K_0 \to \infty$. These are probably not the best potentials for performing the dragging scheme. Spring constants will be chosen such that barriers of say $10kT$ can be crossed. The conventional mesoscopic potentials are not that soft at all. They are probably too stiff. When applying the dragging scheme with these potentials, acceptance probabilities will be low. The simulation probably will get stuck. Note that in this case, the equilibration of the atomistic system and the redrawing of the Gaussian springs do not help. Because of the looseness of the springs compared to the stiffness of the mesoscopic potentials, the factor given by eq 21 remains small.

A solution might be to use softer mesoscopic potentials. The inclusion of a finite K_0 softens the coarse-grained potential $U_{\text{cg}}^{(K_0)}$. This is the potential that should be approximated by the potential, $U_{\text{approx,cg}}$, used mesoscopic simulation. It is, at least at this moment, not clear what values for the spring constants should be taken for molecular simulations. It would be best to be able to generate mesoscopic potentials corresponding to a chosen K_0. Therefore a dual-resolution simulation method for molecular systems that uses the dragging scheme is best combined with a on-the-fly coarse-graining procedure. In this case K_0 can be chosen optimally for the system at hand, and an appropriate coarse-grained potential can be generated. This is a topic of ongoing research.

Besides the factor that depends on the coarse-grained potential, there is also a second contribution to eq 17 due to the springs. How this factor scales with the dimensionality of the system is at present unclear. This factor needs to be kept in bounds also in the high dimensional case, if the scheme is to be useful to speedup molecular simulations. Whether this is possible remains to be seen and will determine the possible applicability in this domain. The large speedup factors of 80 for the one-dimensional case and 20 for the not fully optimized two-dimensional case are encouraging.

In a preliminary molecular study we investigated the dragging scheme to speedup trans−gauche transitions in butane. Here we found that the correct ensemble is sampled. Also high acceptance ratio's could be found by constructing approximate coarse-grained potentials that are well matched to the spring strengths K_0. However, to obtain these high acceptances, the dragging had to be performed so slowly that no effective speedup was achieved.

In the current paper we focused on crossing barriers. As one of the reviewers did point out, a useful application might lie elsewhere. In the coarse-grained dynamics of X, one takes large steps. This means that when dragging the fine-grain configuration, also large steps are taken because it is dragged along with X. This should be compared to local diffusive motion if the configuration is not dragged. Overall this could constitute a speedup in the sampling of phase space.

To a certain extent, the outlined method is a method in need of a good application. It is not clear yet if this application is the one envisioned in the current paper. We have good hopes that the dragging method, or a method derived from it, can be very useful in the realm of molecular simulations. The reason for these hopes is that the outlined method is very flexible, and many possible variants still need to be investigated. One of the flexibilities that is largely unexplored is the fact that during the dragging (i.e., $0 < \lambda < 1$), the used Hamiltonian $H_{\text{drag}}(\Gamma,\lambda)$ can be freely chosen. For example, one can choose to slowly turn of some of the interactions (and turn them on later) or temporarily raise the temperature. The general recipe given by the second line of eq 17 can then be used to generate the correct distribution at $\lambda = 0$ or 1.

APPENDIX

A. Proof of the Neal Dragging Method. We can write the probability density to sample a path in fine-grain phase space, when the initial state obeys the dual-resolution probability density, as

$$\rho(\Gamma_n\ ,\ ...,\ \Gamma_0, X^*, X) = w(\Gamma_n\ ,\ ...,\ \Gamma_1, X^*|\Gamma_0, X, \{\lambda_i\})\rho_{\text{dual}}(\Gamma_0, X) \tag{31}$$

Because all intermediate steps are Markovian processes, the conditional probability for the path can be decomposed as

$$\begin{aligned} &w(\Gamma_n,\ ...,\ \Gamma_1, X^*|\Gamma_0, X, \{\lambda_i\}) \\ &\quad = \text{acc}(\Gamma_n,\ ...,\ \Gamma_1, X^*|\Gamma_0, X) \\ &\quad \times w(X^*|X)\prod_{i=1}^{n} T(\Gamma_i|\Gamma_{i-1}, X^*, X, \lambda_i) \end{aligned} \tag{32}$$

Here we used the transition probabilities as introduced in eqs 9 and 14.

A detailed balance condition can be used to make sure that Γ_n (when averaging over initial and intermediate states) also obeys ρ_{dual}. This detailed balance condition is

$$\begin{aligned} &w(\Gamma_n,\ ...,\ \Gamma_1, X^*|\Gamma_0, X, \{\lambda_i\})\rho_{\text{dual}}(\Gamma_0, X) \\ &\quad = w(\Gamma_0,\ ...,\ \Gamma_{n-1}, X|\Gamma_n, X^*, \{\lambda_i\})\rho_{\text{dual}}(\Gamma_n, X^*) \end{aligned} \tag{33}$$

Here the states Γ_j are reversed, but the sequence of the λ_i's remains unchanged. Inserting eq 32 into this detailed balance condition gives

$$\begin{aligned} &\frac{\text{acc}(\Gamma_n\ ,\ ...,\ \Gamma_1, X^*|\Gamma_0, X)}{\text{acc}(\Gamma_0\ ,\ ...,\ \Gamma_{n-1}, X|\Gamma_n, X^*)} \\ &\quad = \frac{w(X|X^*)}{w(X^*|X)}\prod_{i=1}^{n}\frac{T(\Gamma_{n-i}|\Gamma_{n-i+1}, X, X^*, \lambda_i)}{T(\Gamma_i|\Gamma_{i-1}, X^*, X, \lambda_i)}\frac{\rho_{\text{dual}}(\Gamma_n, X^*)}{\rho_{\text{dual}}(\Gamma_0, X)} \\ &\quad = \frac{w(X|X^*)}{w(X^*|X)}\prod_{i=1}^{n}\frac{T(\Gamma_{i-1}|\Gamma_i, X, X^*, \lambda_{n-i+1})}{T(\Gamma_i|\Gamma_{i-1}, X^*, X, \lambda_i)}\frac{\rho_{\text{dual}}(\Gamma_n, X^*)}{\rho_{\text{dual}}(\Gamma_0, X)} \end{aligned} \tag{34}$$

The last line is obtained from the previous one by relabeling $i \rightarrow n - i + 1$ for the $T(\cdots|\cdots)$'s in the enumerator. This is not the same as the relation in eq 16. The relation of eq 16 is preferred because by means of eq 14, it allows for the transition of eqs 16 to 17. When making the transition to probability densities for the general relation, the normalization constants do not cancel. We want the normalization constants to cancel because these are unknown.

To obtain eq 16 we require that

$$T(\Gamma'|\Gamma, X^*, X, \lambda_{n-i+1}) = T(\Gamma'|\Gamma, X, X^*, \lambda_i) \tag{35}$$

In the simulation we generate the transitions from one Γ to Γ' using the Hamiltonian eq 10. The coarse-grain states X and X^* and the λ's appear through the spring forces. The requirement eq 35 therefore gives that

$$K(\lambda_{n-i+1}) = K^*(\lambda_i) \tag{36}$$

Since λ only provides a parametrization of K, we can take it to increase linearly from 0 to 1. Taking it in accordance with eq 13, we have $\lambda_{n-i+1} = 1 - \lambda_i$ and thus obtain the requirement eq 11.

B. The Slow Dragging Limit. In this appendix we analyze the dragging scheme for the case that dragging is performed (infinitely) slowly with (infinitesimally) small increments $\Delta\lambda$. In this case we can approximate the ratio of acceptance probabilities defined by eq 17 as

$$\begin{aligned} &\frac{f_{\text{approx, cg}}(X)}{f_{\text{approx, cg}}(X^*)}\prod_{i=0}^{n}\frac{f(\Gamma_i, X^*, X, \lambda_{i+1})}{f(\Gamma_i, X^*, X, \lambda_i)} \\ &\quad = \exp[\beta\Delta H_{\text{approx, cg}}]\prod_{i=0}^{n}\exp\left[-\beta(H_{\text{drag}}(\Gamma_i, \lambda_{i+1}) - H_{\text{drag}}(\Gamma_i, \lambda_i))\right] \\ &\quad \approx \exp[\beta\Delta H_{\text{approx, cg}}]\prod_{i=0}^{n}\exp\left[-\beta\Delta\lambda\frac{\partial H_{\text{drag}}(\Gamma_i, \lambda_i)}{\partial\lambda_i}\right] \end{aligned} \tag{37}$$

Because the dragging occurs so slowly, every Γ_i can be assumed to be statistically independent. For large n, many subsequent Γ_i's are sampled for a small increase in λ; this has an averaging effect:

$$\begin{aligned} &\exp[\beta\Delta H_{\text{approx, cg}}]\prod_{i=0}^{n}\exp\left[-\beta\Delta\lambda\frac{\partial H_{\text{drag}}(\Gamma_i, \lambda_i)}{\partial\lambda_i}\right] \\ &\quad \approx \exp[\beta\Delta H_{\text{approx, cg}}]\prod_{i=0}^{n}\left\langle\exp\left[-\beta\Delta\lambda\frac{\partial H_{\text{drag}}(\Gamma_i, \lambda_i)}{\partial\lambda_i}\right]\right\rangle \\ &\quad \approx \exp[\beta\Delta H_{\text{approx, cg}}]\prod_{i=0}^{n}\left[1 - \beta\Delta\lambda\left\langle\frac{\partial H_{\text{drag}}(\Gamma_i, \lambda_i)}{\partial\lambda_i}\right\rangle\right] \\ &\quad \approx \exp[\beta\Delta H_{\text{approx, cg}}]\prod_{i=0}^{n}\left[1 + \Delta\lambda\frac{\partial \ln Z_{\lambda_i}}{\partial\lambda_i}\right] \\ &\quad \approx \exp[\beta\Delta H_{\text{approx, cg}}]\prod_{i=0}^{n}\exp\left[\Delta\lambda\frac{\partial \ln Z_{\lambda_i}}{\partial\lambda_i}\right] \\ &\quad \approx \exp[\beta\Delta H_{\text{approx, cg}}]\exp\left[\sum_{i=0}^{n}\Delta\lambda\frac{\partial \ln Z_{\lambda_i}}{\partial\lambda_i}\right] \\ &\quad \approx \exp[\beta\Delta H_{\text{approx, cg}}]\exp\left[\int_0^1 \mathrm{d}\lambda\frac{\partial \ln Z_\lambda}{\partial\lambda}\right] \\ &\quad = \exp[\beta\Delta H_{\text{approx, cg}}]\exp[\ln Z_1 - \ln Z_0] \\ &\quad = \exp[\beta\Delta(H_{\text{approx, cg}} - H_{cg}^{(K_0)})] \end{aligned} \tag{38}$$

For every i, the expectation value is according to the canonical equilibrium distribution associated with $H_{\text{drag}}(\Gamma,\lambda_i)$. The corresponding partition sum, Z_λ, can be used to compute this expectation value. The Hamiltonian H_{drag} depends on X and X^* according to eq 10. For $\lambda = 0$, the system is only connected to X only and for $\lambda = 1$ only to X^*. In this the partitions sums are given by eq 7 (up to the prefactor which cancels in the computation above).

The final result indicates that, in the slow dragging limit, acceptance ratio's approach 1 when the approximate Hamiltonian used to generate the dynamics of X, i.e., $H_{\text{approx,cg}}$ is close to the real effective Hamiltonian $H_{\text{cg}}^{(K_0)}$. The most severe approximation used in the derivation above is the 'slowness'. At every instance the equilibrium distribution correspondig to λ needs to be sampled. Clearly we want to use the dragging scheme to speed-up computation and therefore not perform the dragging too slowly; hence in practical computations, the acceptance ratio is expected to deviate significantly from the optimal 100% acceptance case.

■ ASSOCIATED CONTENT

Supporting Information. A thorough analysis of the stuck problem is provided. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: e.a.j.f.peters@tue.nl.

■ ACKNOWLEDGMENT

The authors wish to thank Dr. K. S. Lyakhova for fruitful discussions and feedback on earlier versions of the manuscript.

■ REFERENCES

(1) Neal, R. M. *Taking Bigger Metropolis Steps by Draggging Fast Variables*; Technical Report; University of Toronto: Toronto, Canada, October, 2005; http://arxiv.org/PS_cache/math/pdf/0502/0502099v1.pdf.
(2) Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1998**, *102*, 2569–2577.
(3) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.
(4) Muller-Plathe, F. *ChemPhysChem* **2002**, *3*, 754–769.
(5) Lwin, T. Z.; Luo, R. *J. Chem. Phys.* **2005**, *123*, 10.
(6) Lyman, E.; Ytreberg, F. M.; Zuckerman, D. M. *Phys. Rev. Lett.* **2006**, *96*, 028105.
(7) Liu, P.; Shi, Q.; Lyman, E.; Voth, G. A. *J. Chem. Phys.* **2008**, *129*, 8.
(8) Frenkel, D.; Smit, B. *Understanding Molecular Simulation (From Algorithms to Applications)*; Academic Press: San Diego, CA, 2002.
(9) Liu, P.; Voth, G. A. *J. Chem. Phys.* **2007**, *126*, 6.
(10) Lyman, E.; Zuckerman, D. M. *J. Chem. Theory Comput.* **2006**, *2*, 656–666.
(11) Reith, D.; Putz, M.; Muller-Plathe, F. *J. Comput. Chem.* **2003**, *24*, 1624–1636.
(12) Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2005**, *109*, 2469–2473.
(13) Gardiner, C. W. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*; Springer: Berlin, Germany, 1997.
(14) Öttinger, H. *Stochastic Processes in Polymeric Fluids*; Springer Verlag: Berlin, Germany, 1996.