

## Biased Brownian Dynamics for Rate Constant Calculation

Gang Zou,<sup>\*†</sup> Robert D. Skeel,<sup>\*†</sup> and Shankar Subramaniam<sup>\*‡</sup>

<sup>\*</sup>Beckman Institute, <sup>†</sup>Department of Computer Science, and <sup>‡</sup>Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801 USA

**ABSTRACT** An enhanced sampling method—biased Brownian dynamics—is developed for the calculation of diffusion-limited biomolecular association reaction rates with high energy or entropy barriers. Biased Brownian dynamics introduces a biasing force in addition to the electrostatic force between the reactants, and it associates a probability weight with each trajectory. A simulation loses weight when movement is along the biasing force and gains weight when movement is against the biasing force. The sampling of trajectories is then biased, but the sampling is unbiased when the trajectory outcomes are multiplied by their weights. With a suitable choice of the biasing force, more reacted trajectories are sampled. As a consequence, the variance of the estimate is reduced. In our test case, biased Brownian dynamics gives a sevenfold improvement in central processing unit (CPU) time with the choice of a simple centripetal biasing force.

### INTRODUCTION

Many biochemical reactions are classified as diffusion-limited, which means the encounter of the reactants is the most time-limiting part of the reaction. A number of reactions, such as enzyme–substrate and antibody–antigen association, belong to this category. Understanding the nature of the encounter in molecular detail will provide insights into underlying mechanisms of association. Brownian dynamics (BD) simulations have been widely used to calculate the rates of the diffusion-limited reactions. BD theory was developed by Einstein in 1905, Smoluchowski in 1907, 1913, 1915, Fokker in 1913, 1914, and Planck in 1917. The early work in BD is reviewed in an excellent paper by Chandrasekhar (1943). Kramers in 1940 first used BD to study the diffusion-limited reactions. There is a good review of reaction-rate theory by Hänggi et al. (1990). Computer simulations of diffusional motion in biomolecules were first developed by Ermak and McCammon (1978) and Northrup et al. (1984). A state-of-the-art BD simulation code for biomolecules, UHBD, which calculates biomolecular rates of association, is provided by McCammon and co-workers (Davis et al., 1991).

BD simulation is a CPU-intensive computation. The accuracy of the result is typically proportional to the inverse square root of the CPU time. To reduce the simulation time, especially for those reactions with high energy or entropy barriers, an enhanced sampling method—weighted ensemble Brownian dynamics (WEBD)—was developed by Huber and Kim (1996). WEBD maintains an ensemble of particles (in configuration space), and associates a probability weight with each particle. All particles carry out Brownian motion without influencing each other. Periodi-

cally, WEBD splits and merges particles according to their position and weight. This procedure enables better sampling in regions inaccessible to standard BD; thus particles are sampled more near to reaction regions and the accuracy of the result is improved.

Here we propose a novel sampling method—biased Brownian dynamics (bias-BD)—for efficient computation of reaction rates. Bias-BD also associates a weight with each particle and changes the weight as the particle moves. Bias-BD introduces a biasing force in addition to the electrostatic force between the reactants. A particle loses weight when moving along the biasing force and gains weight when moving against the biasing force. As a consequence, particles are sampled more near to the reaction site, albeit with less weight. Bias-BD samples more of the important trajectories and reduces the variance of the result. To obtain the same computational precision in our test case, bias-BD needs only one-seventh the CPU time of the unbiased BD. By selecting a suitable biasing force one can control the sampling density in different regions. Bias-BD simulates only one particle at a time. There is no coupling between particles, as in WEBD, and hence bias-BD is embarrassingly parallel.

Variance-reduced methods for stochastic simulations are discussed in detail in Kloeden and Platen (1992, pages 511–528), which are based on Milstein (1988, page 225) and Wagner (1987, 1988). Milstein (1988) introduces an additional drift term in the stochastic differential equation, which is the same as doing a Girsanov transformation to the underlying probability measure. Wagner (1987, 1988) derives unbiased estimators for functional integrals of stochastic processes based on the general principles of Monte Carlo integration. These methods are used by researchers in other areas, such as Melchior and Öttinger (1995, 1996), and Zuckerman and Woolf (1999). Bias-BD is similar to these variance-reduced methods, but differs from them in three aspects. First, bias-BD is deduced from the discretized stochastic differential equation. The weight calculation is based on the random number distributions at each step and

*Received for publication 27 December 1999 and in final form 26 April 2000.*

Address reprint requests to Gang Zou, University of Illinois at Urbana-Champaign, 405 N. Mathews Ave., Urbana, IL 61801. Tel.: 217-244-7472; Fax: 217-244-2909; E-mail: gangzou@uiuc.edu.

© 2000 by the Biophysical Society

0006-3495/00/08/638/08 \$2.00

bias-BD is valid for each discretized step. Second, bias-BD does not require a unified termination time. In rate constant calculation, each trajectory has a different lifetime. The use of variance-reduced methods in Milstein (1988), Wagner (1987, 1988), and Kloeden and Platen (1992) is not straightforward here. Third, our choice of bias force is different.

In the following sections, we first give an overview of the Northrup, Allison, and McCammon (NAM) method (Northrup et al., 1984); however, our overview is based on a paper by Zhou (1990). Then we describe bias-BD and its weight calculation. Finally, we discuss two test cases and show our choice of bias force. The first is a 3D pure diffusion model, which we can compare with the analytical result. The other is the reaction rate between  $O_2^-$  ions and super-oxide-dismutase (SOD). In the test cases we implement bias-BD in UHBD (Davis et al., 1991) and show a sevenfold improvement in CPU time.

## RATE CONSTANT CALCULATION WITH BROWNIAN DYNAMICS

### Rate constant

We consider the reaction between two types of molecules E (enzyme) and S (substrate). The molecules diffuse in solvent and have certain concentrations. A reaction happens when an S molecule is close enough to an E molecule's binding site. We will compute the reaction rate in a unit volume.

We assume the concentrations do not decrease because of reactions. We also assume concentrations are dilute, i.e., an S molecule's movement is affected by at most one E molecule, and is not affected by others of type S. Under these assumptions, we may consider the relative movement between one pair of molecules. Let  $k$  be the reaction rate for a simplified system, which has one E molecule and S moving relative to the E molecule with a unit relative concentration at infinity. If the actual concentration of S at infinity is  $p_S$ , the reaction rate will be  $kp_S$ . If we assume E's concentration to be  $p_E$ , or there are  $p_E$  E molecules in a unit volume, the reaction rate in a unit volume is  $kp_S p_E$ , because each E molecule provides the same reaction rate independently. In the following discussion we will regard the pair E, S as a "particle" in configuration space.

The movement of S is described by a stochastic differential equation—a diffusion limit Langevin equation, or Langevin equation under strong friction—which means the particle has no inertia, or the particle loses its momentum in a very short time (Ermak and McCammon, 1978):

$$d\mathbf{R} = \frac{D(\mathbf{R})}{k_B T} \mathbf{F}(\mathbf{R}) dt + (2D(\mathbf{R}))^{1/2} d\mathbf{W}(t). \quad (1)$$

In Eq. 1,  $\mathbf{R}(t)$  represents the translational, rotational, and internal coordinates of particle S relative to E.  $D(\mathbf{r})$  is the relative diffusion coefficient between E and S as a function

of coordinates in the configuration space.  $\mathbf{F}(\mathbf{r}) = -\nabla U(\mathbf{r})$  is the electrostatic force that E exerts on S. The force vanishes at infinity, and we may let  $U(\mathbf{r}) = 0$  at infinity.  $\mathbf{W}(t)$  is a vector of independent standard Wiener processes (Kloeden and Platen, 1992, page 70). In the following discussion we treat  $\mathbf{r}$  as a 3D vector; however, the results extend to higher dimensions.

The system is equivalently described by the relative concentration  $p(\mathbf{r})$ . In steady state,  $p(\mathbf{r})$  satisfies the steady-state Smoluchowski equation (Chandrasekhar, 1943; Northrup et al., 1984; Zhou, 1990):

$$\nabla \cdot \mathbf{j}(\mathbf{r}) = 0, \quad (2)$$

where

$$\mathbf{j}(\mathbf{r}) \stackrel{\text{def}}{=} -De^{-U(\mathbf{r})/k_B T} \nabla e^{U(\mathbf{r})/k_B T} p(\mathbf{r}). \quad (3)$$

As mentioned before, we assume  $p(\infty) = 1$ . Let the boundary of molecule E be  $\Gamma_E$ .  $\Gamma_E$  is partitioned into two parts: the binding site  $\Gamma_{E1}$  and the remainder  $\Gamma_{E2}$ . The surface  $\Gamma_{E1}$  is an absorption boundary with  $p(\mathbf{r}) = 0$ . The surface  $\Gamma_{E2}$  is a reflection boundary with  $\mathbf{n} \cdot \mathbf{j}(\mathbf{r}) = 0$ , where  $\mathbf{n}$  is the outward normal direction on  $\Gamma_{E2}$ .

The rate constant  $k$  is given by

$$k = -\oint_{\Gamma} \mathbf{n} dS \cdot \mathbf{j}(\mathbf{r}), \quad (4)$$

where  $\Gamma$  is any surface that encloses  $\Gamma_E$  (Zhou, 1990).

### Domain truncation

The infinite domain rate constant  $k$  is hard to compute directly. If  $p(\mathbf{r})$  and  $U(\mathbf{r})$  are spherically symmetric for  $r \geq q$  with negligible error where  $r$  is the distance between E and S, the Smoluchowski Eq. 2 reduces to an ordinary differential equation for  $r \geq q$  that can be solved by 1D quadrature. The infinite domain problem is then truncated to a finite domain problem. Let

$$p_f(\mathbf{r}) = \frac{e^{-U(q)/k_B T}}{p(q)} p(\mathbf{r}). \quad (5)$$

We call  $p_f(\mathbf{r})$  the finite domain concentration. It satisfies the same equation as  $p(\mathbf{r})$  and similar boundary conditions:

$$\begin{aligned} \nabla \cdot De^{-U(\mathbf{r})/k_B T} \nabla e^{U(\mathbf{r})/k_B T} p_f(\mathbf{r}) &= 0, \\ p_f(q) e^{U(q)/k_B T} &= 1, \\ p_f(\mathbf{r}) &= 0 (\mathbf{r} \in \Gamma_{E1}), \\ \mathbf{n} \cdot De^{-U(\mathbf{r})/k_B T} \nabla e^{U(\mathbf{r})/k_B T} p_f(\mathbf{r}) &= 0 (\mathbf{r} \in \Gamma_{E2}). \end{aligned} \quad (6)$$

Let the rate constant for  $p_f(\mathbf{r})$  be  $k_f$ . From Eqs. 3–5 we have

$$k_f = \frac{e^{-U(q)/k_B T}}{p(q)} k. \quad (7)$$

In Eq. 4, let  $\Gamma$  be a spherical surface with radius  $r \geq q$ . Due to spherical symmetry we have

$$k = 4\pi r^2 D e^{-U(r)/k_B T} \frac{d}{dr} e^{U(r)/k_B T} p(r). \quad (8)$$

So,

$$\frac{d}{dr} e^{U(r)/k_B T} p(r) = \frac{k e^{U(r)/k_B T}}{4\pi D r^2}. \quad (9)$$

Integrating Eq. 9 from  $q$  to  $+\infty$ , we get

$$1 - p(q) e^{U(q)/k_B T} = k \int_q^{+\infty} dr \frac{e^{U(r)/k_B T}}{4\pi D r^2}. \quad (10)$$

Dividing Eq. 10 by  $k$  on both sides and substituting in Eq. 7, we have

$$\frac{1}{k} = \frac{1}{k_f} + \int_q^{+\infty} dr \frac{e^{U(r)/k_B T}}{4\pi D r^2}. \quad (11)$$

In Eq. 11, we see that to compute the infinite domain rate constant  $k$ , we can compute the finite domain rate constant  $k_f$  and a 1D quadrature. The finite domain problem has boundary condition  $p_f(q) = e^{-U(q)/k_B T}$ . Eq. 11 can be explained in the following manner: the term  $1/k$  is similar to the resistance in an electrical system. Two resistances are serially connected. One resistance is from the  $q$  surface to infinity whose value equals the 1D quadrature, and the other resistance is from the  $q$  surface to the reaction site whose value is  $1/k_f$ . The term similar to the voltage in an electrical system is  $e^{U(r)/k_B T} p(r)$ . To measure the resistance, we apply the voltage 1 at one end and voltage 0 at the other end, and measure the current. Eq. 11 tells us that the resistance of two serially connected resistances is the sum of the two.

## FOP and NAM methods

To get the finite domain rate constant  $k_f$ , one approach is to solve the Smoluchowski equation directly. This can be extremely difficult, since it can be a high-dimensional partial differential equation (PDE) if we consider the rotational and internal degrees of freedom of the molecule. Another way is to use random walk, or BD, simulations. Generally speaking, the random walk method provides a simple way to solve high-dimensional PDEs.

One approach to BD simulations is the flux-over-population (FOP) method. In FOP, we set up a time period  $\tau$  and a thin shell with radii from  $q$  to  $q + \Delta q$ , and compute the expected number of particles in the shell  $n = 4\pi q^2 \Delta q e^{-U(q)/k_B T}$ . At the start of time period  $\tau$ , generate  $n$  particles in the shell randomly. During  $\tau$ , let all particles, those in the shell and those inside the  $q$  surface, do Brownian dynamics. At the end of  $\tau$ , count the reacted particles,

divide by  $\tau$  to get the average reaction rate for this  $\tau$  time period, and destroy all the particles beyond the  $q$  surface. The system reaches steady state after a number of time periods. After that, we calculate a rate for each time period and their mean value.

Alternatively, we can calculate a reaction probability  $P(\Delta q, \tau)$  by noting that particles in FOP method do not affect each other and each particle contributes to the reaction rate independently. A particle is generated randomly in the shell. It is allowed to follow a Brownian trajectory where its position is checked every  $\tau$  time.  $P(\Delta q, \tau)$  is the probability that the particle terminates by reacting instead of escaping out of the  $q$  surface. Then,

$$k_f = nP(\Delta q, \tau)/\tau. \quad (12)$$

There is a restriction on  $\Delta q$  and  $\tau$  in FOP. Note that a particle outside the  $q + \Delta q$  surface has a finite probability of entering the  $q$  surface in  $\tau$  time. We require the probability to be small so that it can be neglected. For example, if  $\Delta q > 4 \cdot (2D\tau)^{1/2}$ , the probability is very small.

There is another method to compute  $k_f$ —the NAM method—proposed by Northrup, Allison, and McCammon in 1984 and further explained by Zhou in 1990.

The NAM method further assumes that  $U(\mathbf{r})$  is spherically symmetric for  $r \geq b$  for some  $b < q$ . Let  $\beta(\mathbf{r})$  be the probability that a particle starting at  $\mathbf{r}$  reacts instead of escaping out of the  $q$  surface, and let  $\gamma(\mathbf{r}) = 1 - \beta(\mathbf{r})$ .  $\gamma(\mathbf{r})$  satisfies the steady-state backward Smoluchowski equation and the given boundary conditions (Zhou, 1990):

$$\begin{aligned} e^{U(\mathbf{r})/k_B T} \nabla \cdot D e^{-U(\mathbf{r})/k_B T} \nabla \gamma(\mathbf{r}) &= 0, \\ \gamma(q) &= 1, \\ \gamma(\mathbf{r}) &= 0 (\mathbf{r} \in \Gamma_{E1}), \\ \mathbf{n} \cdot D e^{-U(\mathbf{r})/k_B T} \nabla \gamma(\mathbf{r}) &= 0 (\mathbf{r} \in \Gamma_{E2}). \end{aligned} \quad (13)$$

Comparing Eq. 6 and Eq. 13, we see that  $\gamma(\mathbf{r})$  and  $e^{U(\mathbf{r})/k_B T} p_f(\mathbf{r})$  satisfy the same PDE and same boundary condition; so

$$\gamma(\mathbf{r}) = e^{U(\mathbf{r})/k_B T} p_f(\mathbf{r}). \quad (14)$$

We have

$$k_f = \oint_{\Gamma} \mathbf{n} dS \cdot D e^{-U(\mathbf{r})/k_B T} \nabla \gamma(\mathbf{r}). \quad (15)$$

Taking  $\Gamma$  to be a sphere with radius  $r \geq b$  and noting that  $U(\mathbf{r})$  is spherically symmetric, we have

$$k_f = D e^{-U(r)/k_B T} r^2 \frac{d}{dr} \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi \gamma(r, \theta, \phi). \quad (16)$$

Let

$$\bar{\gamma}(r) = \frac{1}{4\pi} \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi \gamma(r, \theta, \phi); \quad (17)$$

then

$$\frac{k_f e^{U(r)/k_B T}}{4\pi D r^2} = \frac{d}{dr} \bar{\gamma}(r). \quad (18)$$

Integrating Eq. 18 from  $b$  to  $q$ ,

$$k_f \int_b^q dr \frac{e^{U(r)/k_B T}}{4\pi D r^2} = 1 - \bar{\gamma}(b) = \bar{\beta}(b), \quad (19)$$

or

$$\frac{1}{k_f} = \frac{1}{\bar{\beta}(b)} \int_b^q dr \frac{e^{U(r)/k_B T}}{4\pi D r^2}, \quad (20)$$

where

$$\bar{\beta}(r) = \frac{1}{4\pi} \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi \beta(r, \theta, \phi). \quad (21)$$

Thus, the NAM method connects the finite domain rate constant with a probability that could be calculated by BD simulations. Eq. 20 can be explained in the following manner: two resistances are connected in parallel. One resistance is from the  $b$  surface to the  $q$  surface whose value equals the 1D quadrature, and the other resistance is from the  $b$  surface to the reaction site whose value is unknown. Now we add some voltage at the  $b$  surface, and voltage 0 at both the  $q$  surface and the reaction site. With the value  $\bar{\beta}(b)$  that is the percentage of the current going through the unknown resistance, we can calculate the resistance when two resistances are serially connected. In the following discussion we will use  $\beta$  to represent  $\bar{\beta}(b)$ : the average reaction probability starting from a randomly chosen point on the  $b$  surface.

### Error estimation

If the 1D integrals in Eq. 11 and Eq. 20 are evaluated with negligible error, the only error source is from the computation of  $P(\Delta q, \tau)$  or the  $\beta$  value.

To calculate  $\beta$ , let us start a particle  $S$  from  $b$  surface, simulating its movement until either it reacts or escapes. We then assign  $X_i = 1$  or 0 for this trajectory, depending on whether it reacts or not. We simulate many trajectories and obtain many  $X_i$  values. Then we compute their mean as an

approximation of  $\beta$ ,

$$\beta = E(X) \approx \bar{X} = \frac{1}{N} \sum_{i=1}^N X_i, \quad (22)$$

and estimate  $X$ 's variance:

$$\sigma^2(X) \approx \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2. \quad (23)$$

Let  $\Delta\beta$  be the 90% confidence interval of the estimated  $\beta$ . We have

$$\Delta\beta \approx 1.645 \sigma N^{-1/2}. \quad (24)$$

The relative errors for  $k_f$  and  $k$  are

$$\frac{\Delta k_f}{k_f} = \frac{\Delta\beta}{\beta}, \quad (25)$$

$$\frac{\Delta k}{k} = \frac{\Delta\beta}{\beta} \times \frac{k}{k_f} + O\left(\left(\frac{\Delta\beta}{\beta}\right)^2\right). \quad (26)$$

Typically,  $\beta$  is very small due to the energy and entropy barriers between the molecules. The relative error  $\Delta\beta/\beta$  is large. In order to get a suitable relative error, one has to sample many trajectories.

### BIASED BROWNIAN DYNAMICS

As mentioned, bias-BD associates a weight with a particle, and changes the weight as the particle moves. To calculate  $\beta$ , we assign  $X_i = w_i$  or 0 for each trajectory, depending on whether it reacts or not, where  $w_i$  is the final weight of the particle. We then calculate  $\beta$  and  $\Delta\beta$  with Eqs. 22–24, as usual. The effect of bias-BD is that we have many reacted trajectories with small weights instead of a few reacted trajectories with unit weight. Most of the particles that escape have a weight larger than 1. The average weight of all trajectories is  $\sim 1$ . Bias-BD samples more near the binding site and less away from binding site. Thus it reduces the variance of the estimate.

Before examining bias-BD, we examine how standard BD does sampling. By discretizing the diffusion limit Langevin Eq. 1 with the Euler-Maruyama method, introduced into biophysics by Ermak and McCammon (1978), we get

$$\mathbf{R}^{n+1} = \mathbf{R}^n + \frac{D}{k_B T} \mathbf{F}(\mathbf{R}^n) \Delta t + (2D\Delta t)^{1/2} \mathbf{Z}^n, \quad (27)$$

where each  $\mathbf{Z}^n$  is a vector of independent standard Gaussian distributed random numbers that comes from the discretization of the Wiener term  $\mathbf{W}(t)$ .  $\mathbf{Z}^n$  has the probability density function (PDF)

$$p(\mathbf{z}) = (2\pi)^{-3/2} \exp(-\mathbf{z} \cdot \mathbf{z}/2). \quad (28)$$



Bias-BD selects random vector  $\mathbf{Z}'^n$  with a different PDF  $p'(\mathbf{z})$  and associates with each vector a weight function  $w(\mathbf{z})$ , which maintains the following equality:

$$p(\mathbf{z}) = p'(\mathbf{z})w(\mathbf{z}). \quad (29)$$

Thus, the weight distribution product preserves Gaussian distribution. The  $w(\mathbf{z})$  here is the weight for one step of integration. When a particle starts from  $b$  surface, its weight is 1. When a particle moves a step, the new weight equals the old weight times  $w(\mathbf{Z}'^n)$ . The final weight is a product  $w(\mathbf{Z}'^1)w(\mathbf{Z}'^2) \cdots w(\mathbf{Z}'^N)$ .

The new PDF  $p'(\mathbf{z})$  should make reaction more likely. To get a systematic way to derive  $p'(\mathbf{z})$ , let us introduce a biasing force  $\mathbf{F}_b(\mathbf{r})$ .  $\mathbf{F}_b(\mathbf{r})$  acts on the particle like the electrostatic force  $\mathbf{F}(\mathbf{r})$ . The one step integral changes to

$$\begin{aligned} \mathbf{R}^{n+1} = \mathbf{R}^n + \frac{D}{k_B T} (\mathbf{F}(\mathbf{R}^n) + \mathbf{F}_b(\mathbf{R}^n)) \Delta t \\ + (2D\Delta t)^{1/2} \mathbf{Z}'^n, \end{aligned} \quad (30)$$

or,

$$\mathbf{R}^{n+1} = \mathbf{R}^n + \frac{D}{k_B T} \mathbf{F}(\mathbf{R}^n) \Delta t + (2D\Delta t)^{1/2} \mathbf{Z}'^n, \quad (31)$$

where  $\mathbf{Z}'^n = \mathbf{Z}^n + \mathbf{z}_0$ ,  $\mathbf{z}_0 = \mathbf{F}_b(\mathbf{R}^n)(D\Delta t/2)^{1/2}/(k_B T)$ . We see that Eq. 31 is the same as Eq. 27, except it uses random vector  $\mathbf{Z}'^n$  instead of  $\mathbf{Z}^n$ . We know  $\mathbf{Z}^n$  is Gaussian-distributed with PDF  $p(\mathbf{z})$ .  $\mathbf{Z}'^n$  has PDF  $p'(\mathbf{z}) = p(\mathbf{z} - \mathbf{z}_0)$ . So, the weight function is  $w(\mathbf{z}) = p(\mathbf{z})/p(\mathbf{z} - \mathbf{z}_0)$ .

The final part is the determination of the biasing force. The biasing force affects the sampling density significantly. It acts like an additional potential. For instance, we can select this potential artificially to cancel the electrostatic force completely, so we will have equal sampling in all the configuration space. We can add a centripetal biasing force, so we get more sampling near the binding site. There are a multitude of choices.

## TEST CASES

### 3D pure diffusion

We calculate the finite domain rate constant  $k_f$  of a 3D pure diffusion problem as described by the SDE Eq. 1 or by the PDE Eq. 6 with  $U(\mathbf{r}) = 0$  everywhere,  $k_B T = 1$ ,  $D = 1$ ,  $q = 15$ , and reaction occurring at  $r = a = 1$ . The analytical value of the rate constant is  $k_f = 4\pi D/(1/a - 1/q) = 13.464$ .

The 3D pure diffusion can be transformed to a 1D diffusion with an external force. Due to spherical symmetry of  $p_f(\mathbf{r})$ , Eq. 6 can be transformed to

$$\frac{1}{r^2} \frac{d}{dr} r^2 D \frac{d}{dr} p_f(r) = 0. \quad (32)$$

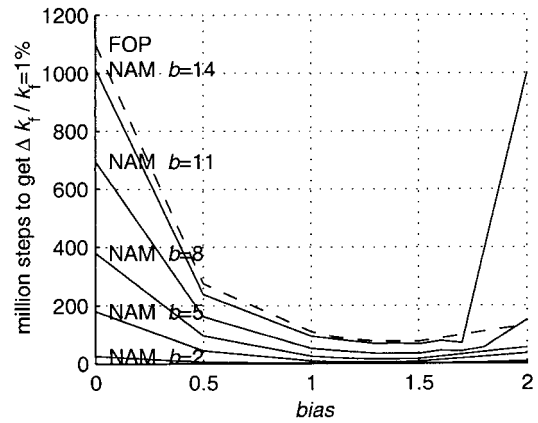


FIGURE 1 Measurement in million integration steps needed to achieve  $\Delta k_f/k_f < 1\%$  in 3D pure diffusion test.

Letting  $p_{1D}(r) = 4\pi r^2 p_f(r)$  and  $U_{1D}(r) = -2k_B T \ln r$ , we have

$$\frac{d}{dr} D e^{-U_{1D}(r)/k_B T} \frac{d}{dr} e^{U_{1D}(r)/k_B T} p_{1D}(r) = 0, \quad (33)$$

which is the 1D Smoluchowski equation with external force  $-(d/dr)U_{1D}(r) = 2k_B T/r$ .

If the bias force cancels the external force in 1D diffusion, we expect to obtain better results. These experiments use the centripetal bias force  $\mathbf{F}_b(\mathbf{r}) = -\text{bias} \times 2k_B T \mathbf{r}/r^2$ , where bias ranges from 0 to 2. When bias = 0, it is standard BD; when bias = 1, the biased system is equivalent to a 1D pure diffusion. We try both FOP and NAM methods with the different strength of bias force. The NAM method is used with various choices for  $b$ . Numerical integration uses variable time steps. In the NAM method, time steps become smaller as  $r$  approaches  $a$  or  $q$ . In the FOP method, time steps become smaller as  $r$  approaches  $a$  and are fixed to the

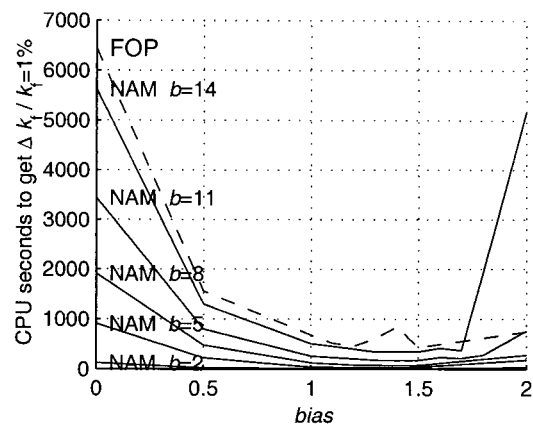


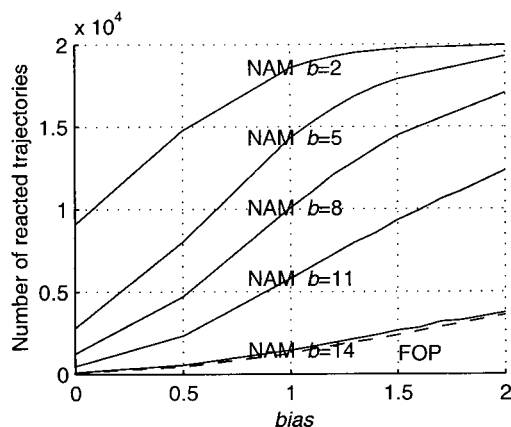
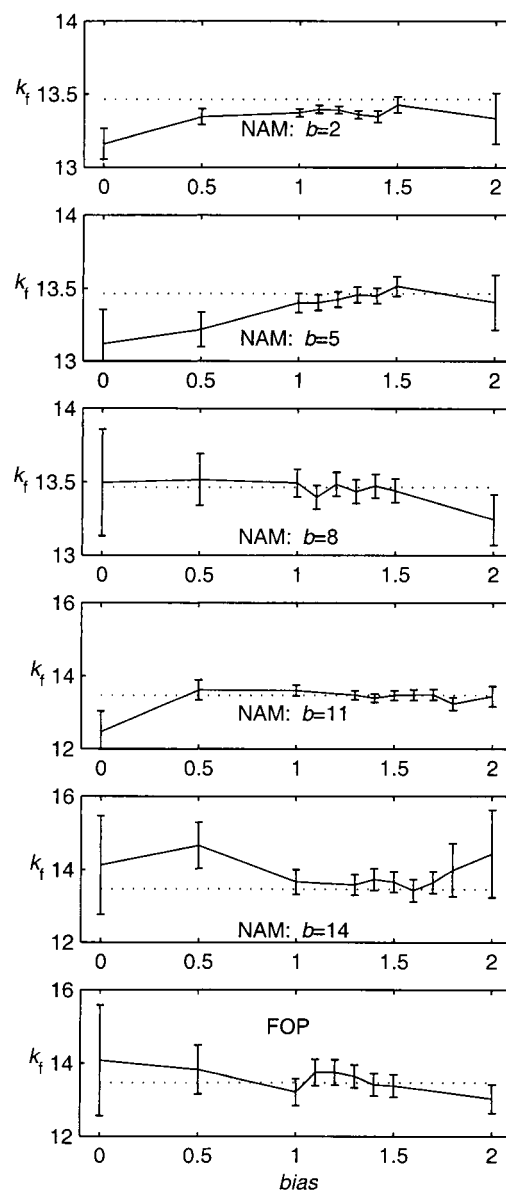
FIGURE 2 Measurement in CPU seconds needed to achieve  $\Delta k_f/k_f < 1\%$  in 3D pure diffusion test.

**TABLE 1** Best bias value and the speedup as measured by number of integration steps in 3D diffusion test

	Best bias	Speedup vs. standard BD
NAM $b = 2$	1.2	76
NAM $b = 5$	1.3	40
NAM $b = 8$	1.3	25
NAM $b = 11$	1.4	21
NAM $b = 14$	1.5	15
FOP	1.5	14

time period  $\tau$  as  $r$  approaches  $q$ . We calculate  $k_f$  with Eq. 12 for the FOP method and Eq. 20 for the NAM method. To measure computational cost, an estimate was calculated of the number of trials required to achieve a relative error of 1% with 90% confidence in the estimated value of  $k_f$ . This is multiplied by the average number of time steps per trial to yield one cost measure, as in Fig. 1, and by seconds of CPU time per trial to give a second cost measure, as in Fig. 2. Table 1 shows the best choice of bias in each configuration and the speedup relative to the standard unbiased BD when measured in number of time steps. Fig. 3 plots the number of reacted trajectories for different choices of bias. Fig. 4 compares the computed  $k_f$  and  $\Delta k_f$  values with the analytical  $k_f$  value.

We see that the FOP method has the same efficiency as the NAM method when the  $b$  surface is very close to the  $q$  surface. The closer the  $b$  surface is to the reaction site, the more efficient is the NAM method. But for a general problem, another error becomes significant for the small  $b$  surface: the asymmetry of the potential  $U(\mathbf{r})$ . Hence, there is a tradeoff in the selection of  $b$ . The best bias value is between 1.2 and 1.5. This implies that the biasing force should be somewhat more centripetal than what is needed to get pure 1D diffusion.

**FIGURE 3** Number of reacted trajectories growing with bias in 3D diffusion test.**FIGURE 4** Comparison of simulated  $k_f$  value and analytical value in 3D diffusion test.

### Rates between $\text{O}_2^-$ ions and SOD

In this experiment we calculate the reaction rate between  $\text{O}_2^-$  ions and SOD. Data of SOD are taken from the Protein Data Bank and the partial charge table is taken from the standard CHARMM force field (Brooks et al., 1983) incorporated into UHBD. SOD is represented by a set of point charges in ionized solvent and fixed in space. The electrostatic potential surrounding SOD is computed by solving the Poisson-Boltzmann equation on a 3D grid.  $\text{O}_2^-$  is treated as a point charge and does BD movements relative to SOD under the influence of the electrostatic field. Reaction is said to happen when  $\text{O}_2^-$  is within 7 Å of the copper atom of

SOD.  $O_2^-$  and each atom of SOD has a hard sphere radius, which prevents  $O_2^-$  from penetrating into the SOD. When  $O_2^-$  has a hard sphere contact with the nonreacting region of SOD, one simply retries the integration step with another Gaussian-distributed random vector.

We use  $F_b(\mathbf{r}) = -2k_B T(\mathbf{r} - \mathbf{r}_0)/|\mathbf{r} - \mathbf{r}_0|^2$  as the biasing force, where  $\mathbf{r}_0$  is the position of the copper atom of SOD. The effect of this force is to have unbiased sampling in the radial direction if there is no electrostatic force. The biasing force and the weight calculation are implemented in UHBD (Davis et al., 1991). Numerical integration uses variable time steps that become smaller as reactants are closer.

Table 2 shows the results. Though the average cost per trajectory of the bias-BD is a little higher than that of the standard BD, the 90% confidence interval  $\Delta\beta$  of the bias-BD is much less. The last line in the table shows that the overall cost of the bias-BD to get 1% relative error is about 1/7 of that of the standard BD.

Fig. 5 compares fluctuations of the  $\beta$  value of bias-BD and standard BD. The solid line is that of bias-BD, and the dashed line is that of standard BD. Each point in the figure represents an estimated  $\beta$  value from 100 trajectories. The horizontal axis shows there are 100 independent runs. We note that bias-BD yields much less fluctuation of the  $\beta$  value, i.e., reduces variance.

## DISCUSSION

In previous sections we reviewed the theory for reaction rate computation and proposed bias-BD to do enhanced BD sampling. We see that the reaction system is similar to an electric system, where the conductance is the rate constant, the voltage is  $e^{U(\mathbf{r})/k_B T} p(\mathbf{r})$ , and the current density is the flux. Bias-BD greatly reduces variance in BD simulations of reactions with high energy or entropy barriers if we select the random variable  $X_i$  to be  $w_i$  or 0, depending on whether it reacts or not, where  $w_i$  is the final weight of the particle.

The main limitation is the ignorance of the choice of bias force that would be optimal for minimal variance. Our choice that the bias force renders the 3D diffusion into a 1D diffusion problem without an external force while being a good choice is perhaps not the best. Our experiments suggest that the bias force could be more centripetal than the force needed to make the 3D diffusion into a 1D diffusion

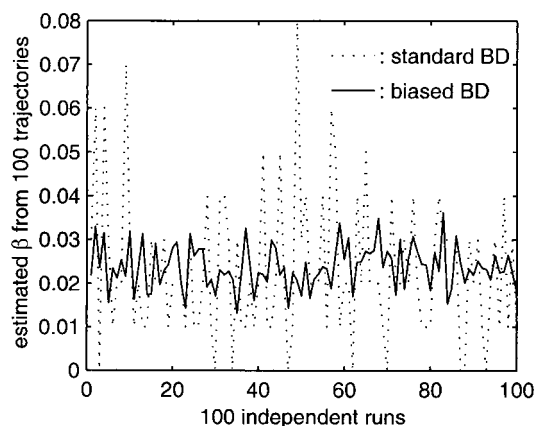


FIGURE 5 Fluctuation of  $\beta$  in bias-BD and standard BD in SOD test.

problem. Further study is in progress to find better choices of biasing forces.

The advantage of BD occurs primarily for high-dimensional configuration spaces. For such problems the entropy barrier will be higher, and the reaction probability will be less. We expect that biased BD can give better improvement. It remains a challenge to find suitable biasing forces in high dimensions that can provide steering automatically.

This work was supported in part by National Science Foundation Grants DBI-9604223, DMS-9600088, MCB-9873199, and DBI-9974555. Also, Gang Zou was supported by a Beckman Institute Research Assistantship and a Computational Science and Engineering Fellowship (University of Illinois).

## REFERENCES

- Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. 1983. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4:187–217.
- Chandrasekhar, S. 1943. Stochastic problem in physics and astronomy. *Rev. Modern Phys.* 15:1–82.
- Davis, M. E., J. D. Madura, B. A. Luty, and J. A. McCammon. 1991. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian dynamics program. *Computer Phys. Comm.* 62:187–197.
- Ermak, D. L., and J. A. McCammon. 1978. Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.* 69:1352–1360.
- Hänggi, P., P. Talkner, and M. Borkovec. 1990. Reaction-rate theory: fifty years after Kramers. *Rev. Modern Phys.* 62:251–342.
- Huber, G. A., and S. Kim. 1996. Weighted-ensemble Brownian dynamics simulations for protein association reactions. *Biophys. J.* 70:97–110.
- Kloeden, P. E., and E. Platen. 1992. Numerical Solution of Stochastic Differential Equations. Springer-Verlag, New York. Second corrected printing, 1995.
- Melchior, M., and H. Öttinger. 1995. Variance reduced simulations of stochastic differential equations. *J. Chem. Phys.* 103:9506–9509.
- Melchior, M., and H. Öttinger. 1996. Variance reduced simulations of polymer dynamics. *J. Chem. Phys.* 105:3316–3331.

TABLE 2 A sevenfold improvement in CPU time in SOD test

10,000 Trajectories	Standard BD	bias-BD
$\beta$	0.02310	0.02369
reacting trajectories	231	6843
$\Delta\beta$	0.00250	0.00085
CPU/trajectory (s)	2.760	3.857
To get $\Delta\beta/\beta < 1\%$ , need trajectories (million)	1.171	0.127
CPU (h)	898	136

- Milstein, G. N. 1988. *The Numerical Integration of Stochastic Differential Equations*. Urals Univ. Press, Sverdlovsk. English ed. by Kluwer Academic Publishers, Dordrecht, The Netherlands, 1995.
- Northrup, S. H., S. A. Allison, and J. A. McCammon. 1984. Brownian dynamics simulation of diffusion-influenced bimolecular reactions. *J. Chem. Phys.* 80:1517–1524.
- Wagner, W. 1987. Unbiased Monte Carlo evaluation of certain functional integrals. *J. Comput. Phys.* 71:21–33.
- Wagner, W. 1988. Monte Carlo evaluation of functionals of solutions of stochastic differential equations, variance reduction and numerical examples. *Stoch. Anal. Appl.* 6:447–468.
- Zhou, H.-X. 1990. On the calculation of diffusive reaction rates using Brownian dynamics simulations. *J. Phys. Chem.* 92:3092–3095.
- Zuckerman, D. M., and T. B. Woolf. 1999. Dynamic reaction paths and rates through importance-sampled stochastic dynamics. *J. Chem. Phys.* 111:9475–9484.