AIMS Mathematics, 9(8): 20793–20810.
DOI: 10.3934/math.20241011
Received: 26 April 2024
Revised: 19 June 2024
Accepted: 21 June 2024
Published: 27 June 2024

https://www.aimspress.com/journal/Math

Research article

On random force correction for large time steps in semi-implicitly discretized overdamped Langevin equations

Takumi Washio[1,2,*], Akihiro Fujii[3] and Toshiaki Hisada[1]

[1] UT-Heart Inc., 178-4-4 Wakashiba, Kashiwa 277-0871, Japan
[2] Graduate School of Frontier Sciences, The University of Tokyo, 178-4-4 Wakashiba, Kashiwa 277-0871, Japan
[3] School of Informatics, Kogakuin University, Shinjuku, Tokyo 163–8677, Japan

* **Correspondence:** Email: washio@ut-heart.com; Tel: +81471355582; Fax: +81471355582.

Abstract: In this study, we focused on the treatment of random forces in a semi-implicitly discretized overdamped Langevin (OL) equation with large time steps. In the usual implicit approach for a nonstochastic mechanical equation, the product of the time interval and Hessian matrix was added to the friction matrix to construct the coefficient matrix for solution updates, which were performed using Newton iteration. When large time steps were used, the additional term, which could be regarded as an artificial friction term, prevented the amplification of oscillations associated with large eigenvalues of the Hessian matrix. In this case, the damping of the high-frequency terms did not cause any discrepancy because they were outside of our interest. However, in OL equations, the friction coefficient was coupled to the random force; therefore, excessive artificial friction may have obscured the effects caused by the stochastic properties of the fluctuations. Consequently, we modified the random force in the proposed semi-implicit scheme so that the total random force was consistent with the friction including the additional artificial term. By deriving a discrete Fokker-Planck (FP) equation from the discretized OL equation, we showed how our modification improved the distribution of the numerical solutions of discrete stochastic processes. Finally, we confirmed the validity of our approach in numerical simulations of a freely jointed chain.

Keywords: overdamped Langevin equation; semi-implicit scheme; random force; friction coefficient; Fokker-Planck equation; Hessian matrix; entropic elasticity
Mathematics Subject Classification: 37H05, 47J25, 60H35, 82C31

1. Introduction

The limited time step size in molecular dynamics (MD) simulations is a major barrier to simulating long-time behaviors of various biological processes, even if coarse-grained molecular models are applied. The limitation originates from the large stiffnesses of strong bonds. In most cases, these strong interactions are necessary to maintain the basic structure of molecules, but the associated small fluctuations do not seem to be essential for molecular deformations of interest. In continuum mechanics, we can overcome the barrier with the use of implicit time integration schemes [1], where the Hessian matrix of the potential function is used to prevent the amplification of oscillations associated with its large eigenvalues. However, the complicated landscapes of MD potential functions hamper the use of implicit schemes because Newton iteration is unstable [2]. In our previous work [3], we introduced a semi-implicit Hessian correction scheme using only the positive parts of Hessian matrices of elemental potentials to avoid instabilities. This scheme allowed us to use time steps 50–200 times larger than those in the explicit Euler-Maruyama (EM) time integration scheme [4]. However, further increases in the time step created considerable errors in the conformational distributions even though the molecular structures were not destroyed.

In this paper, we focus on the damping effect of the Hessian matrix on the fluctuations caused by random forces. In the overdamped Langevin (OL) equation, the increment of the random variable $\boldsymbol{X}$ is given by

$$\boldsymbol{G}d\boldsymbol{X} = \boldsymbol{F}(\boldsymbol{X})dt + \boldsymbol{\Xi}dt, \tag{1.1}$$

where $\boldsymbol{G}$ is a positive-definite matrix representing the friction, $\boldsymbol{F}(\boldsymbol{X})$ is the force as a function of $\boldsymbol{X}$, and $\boldsymbol{\Xi}$ is the random force satisfying $\langle \boldsymbol{\Xi}dt \rangle = \boldsymbol{0}$. We use differential forms and their discrete approximations according to the Itô calculus, as in (1.1), throughout this article. The increment of the probability density function is determined by the averages:

$$\begin{cases} \langle d\boldsymbol{X} \rangle = \boldsymbol{G}^{-1}\boldsymbol{F}(\boldsymbol{X})dt, \\ \langle d\boldsymbol{X} \otimes d\boldsymbol{X} \rangle = \boldsymbol{G}^{-1}\langle \boldsymbol{\Xi}dt \otimes \boldsymbol{\Xi}dt \rangle \boldsymbol{G}^{-1}. \end{cases} \tag{1.2}$$

Here, the first equation represents advection by the force $\boldsymbol{F}(\boldsymbol{X})$. The second equation represents the diffusion of the probability density function. Therefore, the random force is the source of diffusion, which is damped by the friction. In the explicit EM scheme, (1.1) is directly discretized for a finite time interval h as

$$\boldsymbol{G}\Delta\boldsymbol{X} = \boldsymbol{F}\big(\boldsymbol{X}(t)\big)h + \boldsymbol{\Xi}_h h \tag{1.3}$$

to obtain the increment $\Delta\boldsymbol{X} = \boldsymbol{X}(t+h) - \boldsymbol{X}(h)$. Therefore, the impact of the friction coefficient $\boldsymbol{G}$ on the fluctuation of the numerical solution is estimated by replacing dt with h in (1.2). If $\boldsymbol{F}(\boldsymbol{X} + \Delta\boldsymbol{X})$ can be approximated with a positive-semidefinite matrix $\boldsymbol{K}(\boldsymbol{X})$ as

$$\boldsymbol{F}(\boldsymbol{X} + \Delta\boldsymbol{X}) \sim \boldsymbol{F}(\boldsymbol{X}) - \boldsymbol{K}(\boldsymbol{X})\Delta\boldsymbol{X}, \tag{1.4}$$

the following semi-implicit scheme may be applied to enable the use of a large time interval h:

$$\Big(\boldsymbol{G} + h\boldsymbol{K}\big(\boldsymbol{X}(t)\big)\Big)\Delta\boldsymbol{X} = \boldsymbol{F}\big(\boldsymbol{X}(t)\big)h + \boldsymbol{\Xi}_h h. \tag{1.5}$$

In this case, the variance is given by

$$\langle \Delta \boldsymbol{X} \otimes \Delta \boldsymbol{X} \rangle = (\boldsymbol{G} + h\boldsymbol{K})^{-1} \langle \boldsymbol{\Xi}_h h \otimes \boldsymbol{\Xi}_h h \rangle (\boldsymbol{G} + h\boldsymbol{K})^{-1}. \tag{1.6}$$

If there are eigenvalues of $h\boldsymbol{K}$ that are much greater than those of $\boldsymbol{G}$, the damping by $h\boldsymbol{K}$ may underestimate the diffusion. In this study, we pursue a correction scheme for the random force in the semi-implicit EM scheme to shift the diffusion toward the desired direction and magnitude. Comparing (1.3) and (1.5), the coefficient $\boldsymbol{G} + h\boldsymbol{K}\big(\boldsymbol{X}(t)\big)$ on the left-hand side of (1.5) can be regarded as a modified friction coefficient. Therefore, a natural strategy would appear to involve correcting the random force $\boldsymbol{\Xi}_h$ on the right-hand side to be consistent with the modified friction.

We begin our analysis with a linear overdamped Langevin (OL) equation with a positive semidefinite matrix $\boldsymbol{K}$. Through comparison of the numerical solution obtained by the implicit scheme with the analytical solution, we find an appropriate correction to the random force of the form $h\boldsymbol{K}$, where h is the time interval of the temporal discretization. Next, we extend the idea of random force correction to nonlinear OL equations using the Hessian matrix instead of $\boldsymbol{K}$ in the linear case. We derive a discrete Fokker-Planck (FP) equation from the semi-implicitly discretized OL equation with a random force correction. Here, we see that the correction of the random force corresponds to the addition of artificial friction associated with the Hessian matrix. Finally, we validate the proposed semi-implicit scheme on the entropic elasticity problem of the freely jointed chain model.

Grønbech-jensen and Doniach [5] proposed a semi-implicit scheme for the OL equation in which strong bonded interactions are constrained. Though their approach is limited to infinitely stiff bonds, our approach can analyze the effects of the magnitude of strong bonded interactions on the stochastic behavior of molecules, as shown in our numerical experiments. Sweet et al. [6] applied a friction matrix constructed by normal mode partitioning of a mass reweighted Hessian matrix for the underdamped Langevin (UL) equation. Their scheme requires the computation of all eigenvectors of the mass reweighted Hessian matrix to perform the projections on the two subspaces spanned by the low- and high-frequency modes, whereas our scheme requires only the inversion in solving (1.5) and the construction of the additional random forces constructed from the elemental Hessian matrices of the individual strong bonds. This configuration enables faster computations.

2. Preliminary analysis of the linear problem

To determine the essence of the discrepancy of implicit scheme for large times, we start our discussion with a linear OL equation:

$$\gamma \frac{d\boldsymbol{X}(t)}{dt} = -\boldsymbol{K}\boldsymbol{X}(t) + \boldsymbol{\Xi}(t), \tag{2.1}$$

where γ is the friction coefficient, $\boldsymbol{K}$ is a positive semidefinite matrix, and $\boldsymbol{\Xi}(t)$ is the random force vector satisfying

$$\langle \boldsymbol{\Xi}(s)ds \otimes \boldsymbol{\Xi}(u)du \rangle = \delta(s-u) 2k_B T \gamma \boldsymbol{I} ds \tag{2.2}$$

with the Boltzmann constant k_B and the temperature T. Here, $\boldsymbol{I}$ is the identity matrix, and the tensor product $\boldsymbol{a} \otimes \boldsymbol{b}$ of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ can be identified with the square matrix $\boldsymbol{a}\boldsymbol{b}^T$. For a given initial solution $\boldsymbol{X}(t)$ at t, the analytical solution $\boldsymbol{X}(t+h)$ at $t+h$ is as follows:

$$\boldsymbol{X}(t+h) = \exp\left(\frac{h}{\gamma}\boldsymbol{K}\right)^{-1}\left\{\gamma^{-1}\int_0^h \exp\left(\frac{s}{\gamma}\boldsymbol{K}\right)\boldsymbol{\Xi}(t+s)ds + \boldsymbol{X}(t)\right\}. \tag{2.3}$$

Here, the variance of the integrated random force on the right-hand side is calculated using (2.2) to be

$$\begin{aligned}\left\langle\int_0^h \exp\left(\frac{s}{\gamma}\boldsymbol{K}\right)\boldsymbol{\Xi}(t+s)ds \otimes \int_0^h \exp\left(\frac{u}{\gamma}\boldsymbol{K}\right)\boldsymbol{\Xi}(t+u)du\right\rangle &= 2k_BT\gamma\int_0^h \exp\left(2\frac{s}{\gamma}\boldsymbol{K}\right)ds \\ &= k_BT\gamma^2\boldsymbol{K}^{-1}\left(\exp\left(2\frac{h}{\gamma}\boldsymbol{K}\right)-\boldsymbol{I}\right) \\ &= 2k_BT\gamma h\left(\boldsymbol{I}+\frac{h}{\gamma}\boldsymbol{K}+\sum_{l\geq 2}\frac{2^l}{(l+1)!}\left(\frac{h}{\gamma}\boldsymbol{K}\right)^l\right).\end{aligned} \tag{2.4}$$

Here, we assume $x^{-1}(\exp x - 1)=1$ at $x=0$ for the null space of $\boldsymbol{K}$. Let us analyze the relationship between the exact solution at $t+h$ in (2.3) and the temporally discretized solution of the implicit scheme defined as

$$\gamma\frac{\widehat{\boldsymbol{X}}(t+h)-\widehat{\boldsymbol{X}}(t)}{h} = -\boldsymbol{K}\widehat{\boldsymbol{X}}(t+h) + \widehat{\boldsymbol{\Xi}}_h(t), \tag{2.5}$$

where $\widehat{\boldsymbol{\Xi}}_h(t)$ is the discrete random force at t given later. The above equation can be rewritten as

$$\widehat{\boldsymbol{X}}(t+h) = \left(\boldsymbol{I}+\frac{h}{\gamma}\boldsymbol{K}\right)^{-1}\left\{\gamma^{-1}\widehat{\boldsymbol{\Xi}}_h(t)h + \widehat{\boldsymbol{X}}(t)\right\}. \tag{2.6}$$

Here, the matrix $(\boldsymbol{I}+\gamma^{-1}h\boldsymbol{K})$ on the right-hand side can be regarded as the first order Taylor expansion of $\exp(\gamma^{-1}h\boldsymbol{K})$ in (2.3). Furthermore, if we additionally take the first order Taylor expansion of $\gamma^{-1}h\boldsymbol{K}$ in the last term of (2.4), the following condition to be imposed on the random force $\widehat{\boldsymbol{\Xi}}_h(t)$ is derived:

$$\langle\widehat{\boldsymbol{\Xi}}_h(t)h \otimes \widehat{\boldsymbol{\Xi}}_h(t)h\rangle = 2k_BTh(\gamma\boldsymbol{I}+h\boldsymbol{K}). \tag{2.7}$$

Namely, the addition of $h\boldsymbol{K}$ to the identity matrix $\gamma\boldsymbol{I}$, which is used in the explicit EM scheme, may improve the accuracy of implicit schemes. This suggests that the random force should be corrected using the Hessian matrix in semi-implicit schemes for nonlinear problems as well. In other words, the proposed implicit scheme can be regarded as the EM scheme with the random force satisfies (2.7), where $h\boldsymbol{K}$ plays the role of the artificial friction.

Note that the accuracy of the approximation of (2.4) to (2.7) is ensured only for the low frequency components of small eigenvalues of $\gamma^{-1}h\boldsymbol{K}$. However, we are interested in using time steps much larger than that of the stability condition for explicit schemes, i.e.,

$$\frac{h}{\gamma}\rho(\boldsymbol{K}) \gg 1, \tag{2.8}$$

where $\rho(\boldsymbol{K})$ is the spectral radius of $\boldsymbol{K}$. According to (2.4), we cannot expect high accuracy for the high-frequency components for a single time step. However, if the high frequency components are damped over k steps ($k \gg 1$), the temporal change of the stochastic distribution may be correctly obtained. To see this damping effect, we compare the numerical solution after the k steps to the analytical solution. First, we split the analytical solution at $t+kh$ into the stochastic part $\boldsymbol{X}_S$ and the deterministic part $\boldsymbol{X}_D$, where $\boldsymbol{X} = \boldsymbol{X}_S + \boldsymbol{X}_D$, and the two parts are given by

$$\begin{cases} \boldsymbol{X}_S(t+kh) = \exp\left(\frac{kh}{\gamma}\boldsymbol{K}\right)^{-1}\gamma^{-1}\int_0^{kh}\exp\left(\frac{s}{\gamma}\boldsymbol{K}\right)\boldsymbol{\Xi}(t+s)ds\,, \\ \boldsymbol{X}_D(t+kh) = \exp\left(\frac{kh}{\gamma}\boldsymbol{K}\right)^{-1}\boldsymbol{X}(t). \end{cases} \tag{2.9}$$

Here, the variance of $\boldsymbol{X}_S(t+kh)$ is calculated as

$$\begin{aligned} \langle \boldsymbol{X}_S(t+kh)\otimes\boldsymbol{X}_S(t+kh)\rangle &= k_BT\boldsymbol{K}^{-1}\left(\boldsymbol{I}-\exp\left(-\frac{2kh}{\gamma}\boldsymbol{K}\right)\right) \\ &= \frac{2k_BTkh}{\gamma}\left(\frac{2kh}{\gamma}\boldsymbol{K}\right)^{-1}\left(\boldsymbol{I}-\exp\left(-\frac{2kh}{\gamma}\boldsymbol{K}\right)\right). \end{aligned} \tag{2.10}$$

Here, we assume $x^{-1}(1-\exp(-x))=1$ at $x=0$ for the null space of $\boldsymbol{K}$. Similarly, the numerical solution after k steps is split as $\widehat{\boldsymbol{X}}_k = \widehat{\boldsymbol{X}}_{S,k} + \widehat{\boldsymbol{X}}_{D,k}$, where the stochastic part $\widehat{\boldsymbol{X}}_{S,k}$ and deterministic part $\widehat{\boldsymbol{X}}_{D,k}$ are given by

$$\begin{cases} \widehat{\boldsymbol{X}}_{S,k} = \gamma^{-1}\sum_{l=0}^{k-1}\left(\boldsymbol{I}+\frac{h}{\gamma}\boldsymbol{K}\right)^{-(k-l)}\widehat{\boldsymbol{\Xi}}_{h,l}\,h, \\ \widehat{\boldsymbol{X}}_{D,k} = \left(\boldsymbol{I}+\frac{h}{\gamma}\boldsymbol{K}\right)^{-k}\widehat{\boldsymbol{X}}(t). \end{cases} \tag{2.11}$$

Here, $\widehat{\boldsymbol{\Xi}}_{h,l} = \widehat{\boldsymbol{\Xi}}_h(t+lh)$ is the random force applied at the $(l+1)$-th step. From the decorrelation of these random forces and (2.7), the variance of $\widehat{\boldsymbol{X}}_{S,k}$ is

$$\begin{aligned} \langle \widehat{\boldsymbol{X}}_{S,k}\otimes\widehat{\boldsymbol{X}}_{S,k}\rangle &= \gamma^{-2}\sum_{l=0}^{k-1}\left(1+\frac{h}{\gamma}\boldsymbol{K}\right)^{-(k-l)}\langle\widehat{\boldsymbol{\Xi}}_{h,l}h\otimes\widehat{\boldsymbol{\Xi}}_{h,l}h\rangle\left(\boldsymbol{I}+\frac{h}{\gamma}\boldsymbol{K}\right)^{-(k-l)} \\ &= \frac{2k_BTkh}{\gamma}\frac{1}{k}\sum_{l=0}^{k-1}\left(\boldsymbol{I}+\frac{h}{\gamma}\boldsymbol{K}\right)^{-2(k-l)+1}. \end{aligned} \tag{2.12}$$

Thus, the difference between (2.10) and (2.12) can be evaluated by taking the difference of the following two functions, where $x = \gamma^{-1}h\boldsymbol{K}$:

$$\begin{cases} v_k(x) = \frac{1}{2kx}(1-\exp(-2kx)), \\ \hat{v}_k(x) = \frac{1}{k}\sum_{l=0}^{k-1}(1+x)^{-2(k-l)+1}. \end{cases} \tag{2.13}$$

The two functions and their differences are depicted in Figure 1A and B for $k = 1$, 10, and 100. The damping effects for $x \gg o(1)$ are clearly shown for both the functions and the differences. Similarly, the deterministic parts in (2.9) and (2.11) can be evaluated using the functions defined by

$$\begin{cases} d_k(x) = \exp(-kx)\,, \\ \hat{d}_k(x) = (1+x)^{-k}. \end{cases} \tag{2.14}$$

In Figure 1C and D, the same trends as in the stochastic parts are observed.

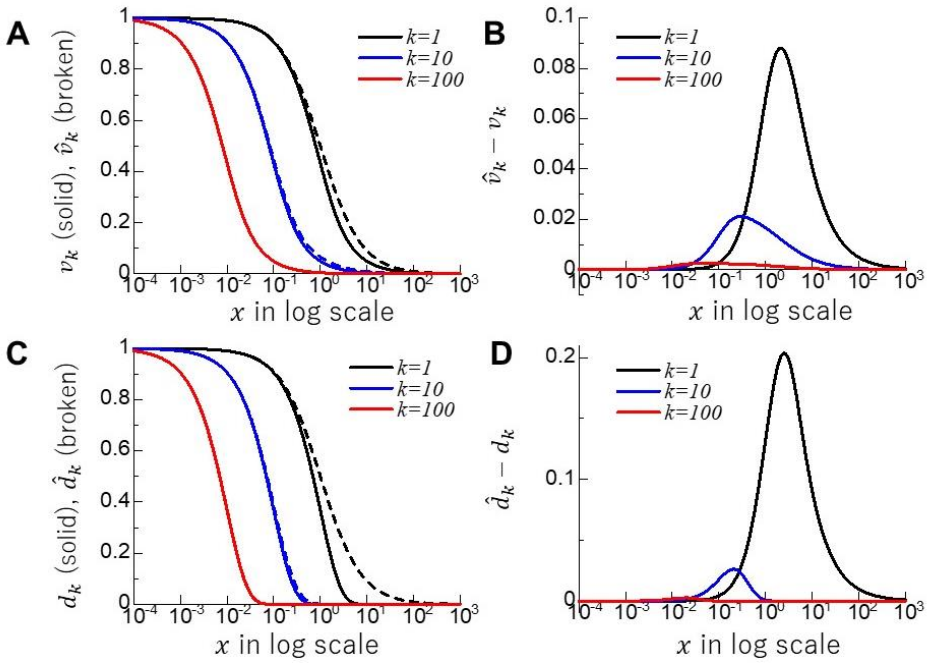

Figure 1. Functions representing the solution updates after k steps of the implicit scheme. The variance functions of the stochastic part are depicted for the analytical solutions (v_k: solid lines) and the discrete implicit solutions ($\hat{v}_k$: broken lines) in **A**. The differences $\hat{v}_k - v_k$ are depicted in **B**. The functions of the deterministic part are depicted for the analytical solutions (d_k: solid lines) and the discrete implicit solutions ($\hat{d}_k$: solid lines) in **C**. The differences $\hat{d}_k - d_k$ are depicted in **D**.

A comparison with the EM scheme

$$\begin{cases} \gamma \dfrac{\breve{\boldsymbol{X}}(t+h) - \breve{\boldsymbol{X}}(t)}{h} = -\boldsymbol{K}\breve{\boldsymbol{X}}(t) + \Xi_h(t), \\ \langle \Xi_h(t)h \otimes \Xi_h(t)h \rangle = 2k_B T\gamma h \boldsymbol{I}, \end{cases} \tag{2.15}$$

will further elucidate the effects of the implicit scheme. In this case, the stochastic and deterministic parts of the numerical solution are represented as

$$\begin{aligned} \breve{\boldsymbol{X}}_{S,k} &= \gamma^{-1} \sum_{l=0}^{k-1} \left(\boldsymbol{I} - \frac{h}{\gamma}\boldsymbol{K} \right)^{k-l} \Xi_{h,l}\, h, \\ \breve{\boldsymbol{X}}_{D,k} &= \left(\boldsymbol{I} - \frac{h}{\gamma}\boldsymbol{K} \right)^{k} \breve{\boldsymbol{X}}(t). \end{aligned} \tag{2.16}$$

The variance of the stochastic part is calculated as

$$\langle \breve{\boldsymbol{X}}_{S,k} \otimes \breve{\boldsymbol{X}}_{S,k} \rangle = \frac{2k_B T k h}{\gamma} \frac{1}{k} \sum_{l=0}^{k-1} \left(\boldsymbol{I} - \frac{h}{\gamma}\boldsymbol{K} \right)^{2(k-l)}. \tag{2.17}$$

In Figure 2, the following functions representing the evolution of solutions are compared with their analytical counterparts for $0 < x < 2$, where convergence is ensured:

$$\begin{cases} \check{v}_k = \gamma^{-1} \sum_{l=0}^{k-1} (1-x)^{2(k-l)}, \\ \check{d}_k = (1-x)^k. \end{cases} \tag{2.18}$$

As in the implicit case, the numerical error around $x = O(1)$ is damped as k increases.

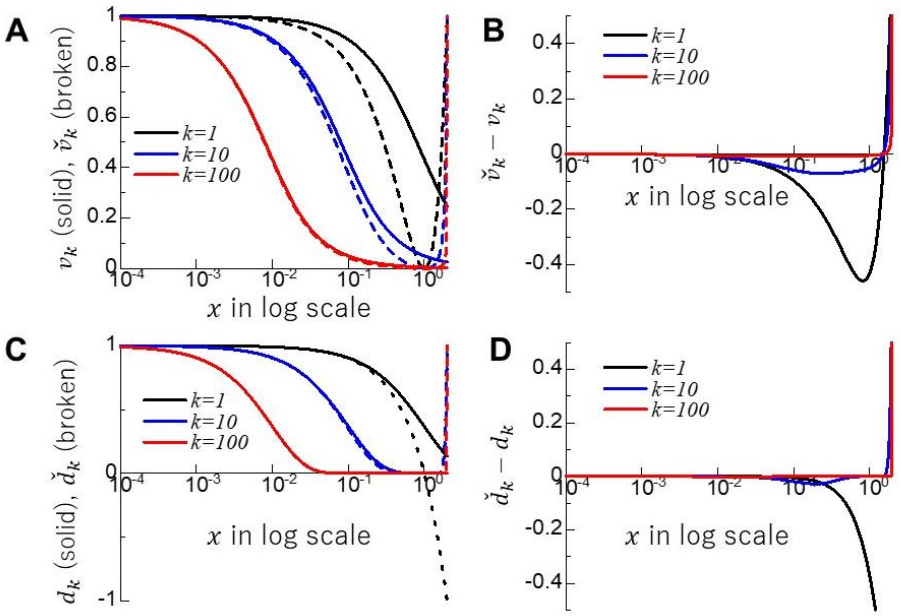

Figure 2. Functions representing the solution updates after k steps of the EM scheme. The variance functions of the stochastic part are depicted for the analytical solutions (v_k: solid lines) and the discrete implicit solutions ($\check{v}_k$: broken lines) in **A**. The differences $\check{v}_k - v_k$ are depicted in **B**. The functions of the deterministic part are depicted for the analytical solutions (d_k: solid lines) and the discrete implicit solutions ($\check{d}_{k_k}$: solid lines) in **C**. The differences $\check{d}_k - d_k$ are depicted in **D**.

3. Modification of friction for nonlinear problems

We consider numerical time integration schemes of the nonlinear OL equation for n particles in three-dimensional space:

$$\boldsymbol{G}\frac{d\boldsymbol{X}(t)}{dt} = -\frac{dU}{d\boldsymbol{X}}(\boldsymbol{X}) + \boldsymbol{\Xi}(t), \tag{3.1}$$

where $\boldsymbol{G}$ is a $3n \times 3n$ positive diagonal matrix representing the friction coefficients imposed on the n particles, $U(\boldsymbol{X})$ is a potential function, and $\boldsymbol{\Xi}(t)$ is a random vector satisfying

$$\langle \boldsymbol{\Xi}(t)dt \otimes \boldsymbol{\Xi}(t)dt \rangle = 2k_B T \boldsymbol{G} dt. \tag{3.2}$$

The potential function consists of bonded and nonbonded interactions, and the thermal effects of the implicit solvent are approximated by the random force. The limitation on the time step size comes from the large eigenvalues of the Hessian matrix $\boldsymbol{H} = \partial^2 U / \partial \boldsymbol{X}^2$. These large eigenvalues originate

from the strong bonded interactions. Because the inversion of a banded matrix consisting of elemental Hessian matrices of the strong bonded interactions is computationally efficient, the use of a semi-implicit method is natural. In the following, we start with the EM scheme, and modify the friction coefficient and the random force to enable a large time interval.

In accordance with the Ito integral for a constant time interval h as in the EM scheme, we must update the particle coordinates $\boldsymbol{X}_h$ from t to $t+h$ as follows:

$$\boldsymbol{G}\frac{\boldsymbol{X}_h(t+h)-\boldsymbol{X}_h(t)}{h}=-\frac{dU}{d\boldsymbol{X}}\big(\boldsymbol{X}_h(t)\big)+\boldsymbol{\Xi}_h(t), \tag{3.3}$$

where the right-hand side involves the force $-dU/d\boldsymbol{X}$ at t and the discrete random force satisfies

$$\langle \boldsymbol{\Xi}_h(t)h\otimes\boldsymbol{\Xi}_h(t)h\rangle=2k_BT\boldsymbol{G}h. \tag{3.4}$$

When applying the above scheme to MD simulations, we encounter a severe restriction of the time interval h owing to the stability condition:

$$\rho\left(\boldsymbol{I}-h\boldsymbol{G}^{-1}\frac{\partial^2U}{\partial\boldsymbol{X}^2}\right)\le 1. \tag{3.5}$$

In our previous work [3], we introduced a semi-implicit Hessian correction scheme (SimHec) with the intention of replacing $-\frac{dU}{d\boldsymbol{X}}\big(\boldsymbol{X}_h(t)\big)$ in (3.3) by a linear approximation of $-\frac{dU}{d\boldsymbol{X}}\big(\boldsymbol{X}_h(t+h)\big)$ as follows:

$$\left(\boldsymbol{G}+h\widetilde{\boldsymbol{H}}\left(\widetilde{\boldsymbol{X}}_h(t)\right)\right)\frac{\widetilde{\boldsymbol{X}}_h(t+h)-\widetilde{\boldsymbol{X}}_h(t)}{h}=-\frac{dU}{d\boldsymbol{X}}\left(\widetilde{\boldsymbol{X}}_h(t)\right)+\boldsymbol{\Xi}_h(t). \tag{3.6}$$

Here, $\widetilde{\boldsymbol{H}}(\boldsymbol{X})$ is a $3n\times 3n$ symmetric matrix designed to approximate the positive part of the Hessian matrix of U at $\boldsymbol{X}$. Based on the derivation of (2.7), we regard the coefficient $h\widetilde{\boldsymbol{H}}\left(\widetilde{\boldsymbol{X}}_h(t)\right)$ on the left-hand side of (3.6) as the artificial friction and apply the corrected random force as follows:

$$\left(\boldsymbol{G}+h\widetilde{\boldsymbol{H}}\left(\widehat{\boldsymbol{X}}_h(t)\right)\right)\frac{\widehat{\boldsymbol{X}}_h(t+h)-\widehat{\boldsymbol{X}}_h(t)}{h}=-\frac{dU}{d\boldsymbol{X}}\left(\widehat{\boldsymbol{X}}_h(t)\right)+\boldsymbol{\Xi}_h(t)+\widetilde{\boldsymbol{\Xi}}_h(t), \tag{3.7}$$

where the additional random force $\widetilde{\boldsymbol{\Xi}}_h(t)$ satisfies the conditions

$$\begin{cases}\langle \boldsymbol{\Xi}_h(t)h\otimes\widetilde{\boldsymbol{\Xi}}_h(t)h\rangle=\boldsymbol{0},\\ \langle \widetilde{\boldsymbol{\Xi}}_h(t)h\otimes\widetilde{\boldsymbol{\Xi}}_h(t)h\rangle=2k_BTh\widetilde{\boldsymbol{H}}\left(\widehat{\boldsymbol{X}}_h(t)\right)h.\end{cases} \tag{3.8}$$

Here, the first condition indicates that the additional random force $\widetilde{\boldsymbol{\Xi}}_h$ must be uncorrelated with the original random force $\boldsymbol{\Xi}_h(t)$, and the second condition corresponds to the consistency of the additional random force and the artificial friction $h\widetilde{\boldsymbol{H}}$. From these conditions, we can easily derive the consistency condition of the total friction $\widehat{\boldsymbol{G}}_h\left(\widehat{\boldsymbol{X}}_h(t)\right)=\boldsymbol{G}+h\widetilde{\boldsymbol{H}}\left(\widehat{\boldsymbol{X}}_h(t)\right)$ and the total random force $\widehat{\boldsymbol{\Xi}}_h(t)=\boldsymbol{\Xi}_h(t)+\widetilde{\boldsymbol{\Xi}}_h(t)$, which is

$$\langle \widehat{\boldsymbol{\Xi}}_h(t)h\otimes\widehat{\boldsymbol{\Xi}}_h(t)h\rangle=2k_BT\,\widehat{\boldsymbol{G}}_h\left(\widehat{\boldsymbol{X}}_h(t)\right)h. \tag{3.9}$$

In the following, we discuss the meaning of this consistency condition in the context of the probability

density. Hereafter, we refer to this method as the "semi-implicit Hessian correction scheme with random force correction" (SimHec-RC).

4. Impact of random force correction on probability density

As shown in the literature [7–9], the property (3.2) of the random force is key to deriving the FP equation when applying the Ito integral to the OL Eq (3.1). Here, we trace the process from the OL equation to the FP equation based on the discrete temporal integration of SimHec-RC.

In SimHec-RC, the update $\Delta\widehat{\boldsymbol{X}}_h = \widehat{\boldsymbol{X}}_h(t+h) - \widehat{\boldsymbol{X}}_h(t)$ is performed as follows:

$$\Delta\widehat{\boldsymbol{X}}_h = -\widehat{\boldsymbol{G}}_h\left(\widehat{\boldsymbol{X}}_h(t)\right)^{-1}\frac{dU}{d\boldsymbol{X}}\left(\widehat{\boldsymbol{X}}_h(t)\right)h + \widehat{\boldsymbol{G}}_h\left(\widehat{\boldsymbol{X}}_h(t)\right)^{-1}\widehat{\boldsymbol{\Xi}}_h(t)h. \tag{4.1}$$

Therefore, we obtain

$$\begin{aligned}\langle\Delta\widehat{\boldsymbol{X}}_h\rangle &= -\widehat{\boldsymbol{G}}_h\left(\widehat{\boldsymbol{X}}_h(t)\right)^{-1}\frac{dU}{d\boldsymbol{X}}\left(\widehat{\boldsymbol{X}}_h(t)\right)h,\\ \langle\Delta\widehat{\boldsymbol{X}}_h\otimes\Delta\widehat{\boldsymbol{X}}_h\rangle &= 2k_BT\widehat{\boldsymbol{G}}_h\left(\widehat{\boldsymbol{X}}_h(t)\right)^{-1}h + \boldsymbol{O}(h^2).\end{aligned} \tag{4.2}$$

Here, we have derived the second equation assuming (3.9) as follows:

$$\begin{aligned}\langle\widehat{\boldsymbol{G}}_h\left(\widehat{\boldsymbol{X}}_h\right)^{-1}\widehat{\boldsymbol{\Xi}}_h h\otimes\widehat{\boldsymbol{G}}_h\left(\widehat{\boldsymbol{X}}_h\right)^{-1}\widehat{\boldsymbol{\Xi}}_h h\rangle &= \widehat{\boldsymbol{G}}_h\left(\widehat{\boldsymbol{X}}_h\right)^{-1}\langle\widehat{\boldsymbol{\Xi}}_h h\otimes\widehat{\boldsymbol{\Xi}}_h h\rangle\widehat{\boldsymbol{G}}_h\left(\widehat{\boldsymbol{X}}_h\right)^{-1} =\\ &2k_BTh\widehat{\boldsymbol{G}}_h\left(\widehat{\boldsymbol{X}}_h\right)^{-1}.\end{aligned} \tag{4.3}$$

Note that the higher order products of $\widehat{\boldsymbol{G}}_h\left(\widehat{\boldsymbol{X}}_h\right)^{-1}\widehat{\boldsymbol{\Xi}}_h h$ are $\boldsymbol{O}(h^2)$. If we replace the modified random force $\widehat{\boldsymbol{\Xi}}_h(t)$ with the original random force $\boldsymbol{\Xi}_h(t)$ as in the original SimHec in (3.6), the second term is replaced by

$$\langle\Delta\widetilde{\boldsymbol{X}}_h\otimes\Delta\widetilde{\boldsymbol{X}}_h\rangle = 2k_BTh\widehat{\boldsymbol{G}}_h\left(\widetilde{\boldsymbol{X}}_h(t)\right)^{-1}\boldsymbol{G}\widehat{\boldsymbol{G}}_h\left(\widetilde{\boldsymbol{X}}_h(t)\right)^{-1} + \boldsymbol{O}(h^2). \tag{4.4}$$

Here, $\langle\Delta\widetilde{\boldsymbol{X}}_h\otimes\Delta\widetilde{\boldsymbol{X}}_h\rangle$ differs by the factor $\boldsymbol{G}\widehat{\boldsymbol{G}}_h\left(\widetilde{\boldsymbol{X}}_h(t)\right)^{-1}$ from $\langle\Delta\widehat{\boldsymbol{X}}_h\otimes\Delta\widehat{\boldsymbol{X}}_h\rangle$ in (4.2). Below, we evaluate the impact of this difference on the probability density functions produced by these discrete stochastic processes.

From (4.2), we can derive the discrete version of Ito's lemma [10] for any smooth function f using Taylor expansion:

$$\begin{aligned}&\langle f\left(\widehat{\boldsymbol{X}}_h(t)+\Delta\widehat{\boldsymbol{X}}_h\right)\rangle - f(\widehat{\boldsymbol{X}}_h(t)) = \langle f\left(\widehat{\boldsymbol{X}}_h(t)+\Delta\widehat{\boldsymbol{X}}_h\right) - f\left(\widehat{\boldsymbol{X}}_h(t)\right)\rangle\\ &= \langle\frac{df}{d\boldsymbol{X}}\left(\widehat{\boldsymbol{X}}_h(t)\right):\Delta\widehat{\boldsymbol{X}}_h + \frac{1}{2}\frac{d^2f}{d\boldsymbol{X}^2}\left(\widehat{\boldsymbol{X}}_h(t)\right):\Delta\widehat{\boldsymbol{X}}_h\otimes\Delta\widehat{\boldsymbol{X}}_h\rangle + O(h^2)\\ &= h\widehat{\boldsymbol{G}}_h\left(\widehat{\boldsymbol{X}}_h(t)\right)^{-1}:\left(-\frac{df}{d\boldsymbol{X}}\left(\widehat{\boldsymbol{X}}_h(t)\right)\otimes\frac{dU}{d\boldsymbol{X}}\left(\widehat{\boldsymbol{X}}_h(t)\right) + k_BT\frac{d^2f}{d\boldsymbol{X}^2}\left(\widehat{\boldsymbol{X}}_h(t)\right)\right)\\ &+ O(h^2).\end{aligned} \tag{4.5}$$

Here, the average $\langle*\rangle$ is taken over the discrete stochastic processes from t to $t+h$, and the symbol ":" represents the dot product of two vectors or two matrices. Let $\hat{P}_h\left(\widehat{\boldsymbol{X}}_h, t\right)$ be the probability density function of SimHec-RC in (3.7). Then, we have the following relationship between the integrals over the conformation spaces $\widehat{\Omega}_h(t)$ and $\widehat{\Omega}_h(t+h)$ at times t and $t+h$, respectively:

$$\begin{aligned}&\int \langle f\big(\widehat{\boldsymbol{X}}_h(t)+\Delta\widehat{\boldsymbol{X}}_h\big)\rangle \hat{P}_h\big(\widehat{\boldsymbol{X}}_h(t),t\big)d\widehat{\Omega}_h\,(t)\\&=\int f\left(\widehat{\boldsymbol{X}}_h(t+h)\right)\hat{P}_h\big(\widehat{\boldsymbol{X}}_h(t+h),t+h\big)d\widehat{\Omega}_h(t+h).\end{aligned}\tag{4.6}$$

Here, the averaging operation $\langle f\big(\widehat{\boldsymbol{X}}_h+\Delta\widehat{\boldsymbol{X}}_h\big)\rangle$ on the left-hand side is taken over $\Delta\widehat{\boldsymbol{X}}_h$. By using Eq (4.6) followed by Eq (4.5) and applying integration by parts, we obtain the following equations:

$$\begin{aligned}&\int f\big(\widehat{\boldsymbol{X}}_h\big)\frac{1}{h}\Big(\hat{P}_h\big(\widehat{\boldsymbol{X}}_h,t+h\big)-\hat{P}_h\big(\widehat{\boldsymbol{X}}_h,t\big)\Big)\,d\widehat{\Omega}_h\\&=\frac{1}{h}\int\Big(\langle f\big(\widehat{\boldsymbol{X}}_h+\Delta\widehat{\boldsymbol{X}}_h\big)\rangle-f\big(\widehat{\boldsymbol{X}}_h\big)\Big)\hat{P}_h\big(\widehat{\boldsymbol{X}}_h,t\big)\,d\widehat{\Omega}_h\\&=\int\left\{\widehat{\boldsymbol{G}}_h\big(\widehat{\boldsymbol{X}}_h\big)^{-1}:\left(-\frac{df}{d\boldsymbol{X}}\big(\widehat{\boldsymbol{X}}_h\big)\otimes\frac{dU}{d\boldsymbol{X}}\big(\widehat{\boldsymbol{X}}_h\big)+k_BT\frac{d^2f}{d\boldsymbol{X}^2}\big(\widehat{\boldsymbol{X}}_h\big)\right)\right\}\hat{P}_h\big(\widehat{\boldsymbol{X}}_h,t\big)\,d\widehat{\Omega}_h\\&+O(h)\\&=\int f\big(\widehat{\boldsymbol{X}}_h\big)\left\{\frac{d}{d\boldsymbol{X}}:\left(\widehat{\boldsymbol{G}}_h(\boldsymbol{X})^{-1}\frac{dU}{d\boldsymbol{X}}(\boldsymbol{X})\hat{P}_h(\boldsymbol{X},t)\right.\right.\\&\left.\left.+k_BT\frac{d}{d\boldsymbol{X}}:\Big(\widehat{\boldsymbol{G}}_h(\boldsymbol{X})^{-1}\hat{P}_h(\boldsymbol{X},t)\Big)\right)\right|_{\boldsymbol{X}=\widehat{\boldsymbol{X}}_h}\Bigg\}d\widehat{\Omega}_h+O(h).\end{aligned}\tag{4.7}$$

Because the above equation holds for any smooth function f, we obtain

$$\begin{aligned}&\frac{1}{h}\Big(\hat{P}_h(\boldsymbol{X},t+h)-\hat{P}_h(\boldsymbol{X},t)\Big)\\&=\frac{d}{d\boldsymbol{X}}:\Big\{\widehat{\boldsymbol{G}}_h(\boldsymbol{X})^{-1}\frac{dU}{d\boldsymbol{X}}(\boldsymbol{X})\hat{P}_h(\boldsymbol{X},t)+k_BT\frac{d}{d\boldsymbol{X}}:\Big(\widehat{\boldsymbol{G}}_h(\boldsymbol{X})^{-1}\hat{P}_h(\boldsymbol{X},t)\Big)\Big\}+O(h).\end{aligned}\tag{4.8}$$

This is a discrete version of the FP equation in which $\widehat{\boldsymbol{G}}_h(\boldsymbol{X})$ corresponds to the friction coefficients. Similarly, from (4.4), we obtain the following equation for the probability density function $\tilde{P}_h(\boldsymbol{X},t)$ for the stochastic process given by SimHec in (3.6) without the random force correction:

$$\begin{aligned}&\frac{1}{h}\Big(\tilde{P}_h(\boldsymbol{X},t+h)-\tilde{P}_h(\boldsymbol{X},t)\Big)\\&=\frac{d}{d\boldsymbol{X}}:\Big\{\widehat{\boldsymbol{G}}_h(\boldsymbol{X})^{-1}\frac{dU}{d\boldsymbol{X}}(\boldsymbol{X})\tilde{P}_h(\boldsymbol{X},t)+k_BT\frac{d}{d\boldsymbol{X}}:\Big(\widehat{\boldsymbol{G}}_h(\boldsymbol{X})^{-1}\boldsymbol{G}\widehat{\boldsymbol{G}}_h(\boldsymbol{X})^{-1}\tilde{P}_h(\boldsymbol{X},t)\Big)\Big\}\\&+O(h).\end{aligned}\tag{4.9}$$

Here, we see that the dispersion is overdamped by a factor of $\widehat{\boldsymbol{G}}_h\left(\widetilde{\boldsymbol{X}}(t)\right)^{-1}\boldsymbol{G}$ compared with (4.8). The advantage of the random force correction is clear when $\widehat{\boldsymbol{G}}_h$ is a constant matrix. The term inside the brackets on the right-hand side of (4.8) is rewritten as

$$\begin{aligned}&\widehat{\boldsymbol{G}}_h(\boldsymbol{X})^{-1}\frac{dU}{d\boldsymbol{X}}(\boldsymbol{X})\hat{P}_h(\boldsymbol{X},t)+k_BT\frac{d}{d\boldsymbol{X}}:\left(\widehat{\boldsymbol{G}}_h(\boldsymbol{X})^{-1}\hat{P}_h(\boldsymbol{X},t)\right)\\&=\widehat{\boldsymbol{G}}_h(\boldsymbol{X})^{-1}\left\{\frac{dU}{d\boldsymbol{X}}(\boldsymbol{X})\hat{P}_h(\boldsymbol{X},t)+k_BT\frac{d\hat{P}_h}{d\boldsymbol{X}}(\boldsymbol{X},t)\right\}\\&+k_BT\left\{\frac{d}{d\boldsymbol{X}}:\widehat{\boldsymbol{G}}_h{}^{-1}\right\}(\boldsymbol{X})\hat{P}_h(\boldsymbol{X},t).\end{aligned}\tag{4.10}$$

Thus, we see that the probability density of the FP equation converges to the Boltzmann distribution, i.e.,

$$\lim_{t\to\infty}\hat{P}_h(\boldsymbol{X},t)\propto\exp\left(-\frac{U(\boldsymbol{X})}{k_BT}\right).\tag{4.11}$$

Equation (4.10) also suggests that the accuracy of the obtained time transients of the distribution depends on the magnitude of $\left\|\frac{d}{d\boldsymbol{X}}:\boldsymbol{G}_h{}^{-1}\right\|$. Therefore, it may be important to minimize $\left\|\frac{d}{d\boldsymbol{X}}\widetilde{\boldsymbol{H}}\right\|$ in the construction of $\widetilde{\boldsymbol{H}}$. Physically, $\frac{d}{d\boldsymbol{X}}:\boldsymbol{G}_h{}^{-1}\neq\boldsymbol{0}$ means that the thermal effect of the solvent depends on the conformation $\boldsymbol{X}$. Thus, the probability density does not converge to the Boltzmann distribution in thermal equilibrium for a finite h.

5. Combining the artificial friction matrix and random force

Here, we consider combining the $\widetilde{\boldsymbol{H}}(\boldsymbol{X})$ and $\widetilde{\boldsymbol{\Xi}}_h$ that satisfy the conditions in (3.8). In general, the potential function $U(\boldsymbol{X})$ in MD is given as a sum of elemental potentials between two, three, or four particles, as follows:

$$U(\boldsymbol{X})=\sum_k V_k(\boldsymbol{P}_k\boldsymbol{X}),\tag{5.1}$$

where $\boldsymbol{P}_k$ represents projection onto the subspace Ω_k of the subset of particles for which the potential V_k is defined. For every potential V_k, if we can construct $\widetilde{\boldsymbol{H}}_k$ and $\widetilde{\boldsymbol{\Xi}}_{k,h}$ on Ω_k such that the conditions

$$\begin{cases}\langle\boldsymbol{\Xi}_h(t)h\otimes\widetilde{\boldsymbol{\Xi}}_{k,h}(t)h\rangle=\boldsymbol{0},\\\langle\widetilde{\boldsymbol{\Xi}}_{h,k}(t)h\otimes\widetilde{\boldsymbol{\Xi}}_{h,l}(t)h\rangle=\delta_{kl}2k_BTh\widetilde{\boldsymbol{H}}_k\big(\boldsymbol{P}_k\boldsymbol{X}(t)\big)h,\end{cases}\tag{5.2}$$

are fulfilled, then it is straightforward to show that the sums

$$\begin{cases}\widetilde{\boldsymbol{H}}\big(\boldsymbol{X}(t)\big)=\sum_{\boldsymbol{k}}\widetilde{\boldsymbol{H}}_k\big(\boldsymbol{P}_k\boldsymbol{X}(t)\big),\\\widetilde{\boldsymbol{\Xi}}_h(t)=\sum_{\boldsymbol{k}}\widetilde{\boldsymbol{\Xi}}_{k,h}(t),\end{cases}\tag{5.3}$$

satisfy the conditions in (3.8). Here, we identify $\widetilde{\boldsymbol{H}}_k$ and $\widetilde{\boldsymbol{\Xi}}_{k,h}$ on the subspace Ω_k with $\boldsymbol{P}_k{}^T\widetilde{\boldsymbol{H}}_k\boldsymbol{P}_k$ and $\boldsymbol{P}_k{}^T\widetilde{\boldsymbol{\Xi}}_{k,h}$ on the total space, respectively, for the sake of simplicity. We can realize the uncorrelatedness of random forces between different elemental potentials using sequences of uncorrelated random numbers. Therefore, the remaining problems are (i) constructing $\widetilde{\boldsymbol{H}}_k$ for V_k, and (ii) constructing a $\widetilde{\boldsymbol{\Xi}}_{k,h}(t)$ that fulfills the condition:

$$\langle \widetilde{\boldsymbol{\Xi}}_{h,k}(t)h \otimes \widetilde{\boldsymbol{\Xi}}_{h,k}(t)h \rangle = 2k_B T h \widetilde{\boldsymbol{H}}_k\big(\boldsymbol{X}_k(t)\big)h,\ \boldsymbol{X}_k(t) = \boldsymbol{P}_k \boldsymbol{X}(t). \tag{5.4}$$

As an example, we consider the potential $V(\boldsymbol{x}_1, \boldsymbol{x}_2) = \varphi\big(r_{1,2}\big)$, which is a function of the distance $r_{1,2} = \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|$. In this case, the Hessian matrix is

$$\boldsymbol{H}_{1,2} = \begin{bmatrix} \dfrac{\partial^2 V}{\partial {\boldsymbol{x}_1}^2} & \dfrac{\partial^2 V}{\partial \boldsymbol{x}_1 \boldsymbol{x}_2} \\ \dfrac{\partial^2 V}{\partial \boldsymbol{x}_2 \boldsymbol{x}_1} & \dfrac{\partial^2 V}{\partial {\boldsymbol{x}_2}^2} \end{bmatrix} = \ddot{\varphi}\big(r_{1,2}\big)\boldsymbol{A}_{1,2} + \frac{\dot{\varphi}\big(r_{1,2}\big)}{r_{1,2}}\boldsymbol{B}_{1,2}, \tag{5.5}$$

where the matrices $\boldsymbol{A}_{1,2}$ and $\boldsymbol{B}_{1,2}$ are defined as

$$\boldsymbol{A}_{1,2} = \begin{bmatrix} \boldsymbol{n} \otimes \boldsymbol{n} & -\boldsymbol{n} \otimes \boldsymbol{n} \\ -\boldsymbol{n} \otimes \boldsymbol{n} & \boldsymbol{n} \otimes \boldsymbol{n} \end{bmatrix},\ \boldsymbol{B}_{1,2} = \begin{bmatrix} \boldsymbol{I}_3 - \boldsymbol{n} \otimes \boldsymbol{n} & -\boldsymbol{I}_3 + \boldsymbol{n} \otimes \boldsymbol{n} \\ -\boldsymbol{I}_3 + \boldsymbol{n} \otimes \boldsymbol{n} & \boldsymbol{I}_3 - \boldsymbol{n} \otimes \boldsymbol{n} \end{bmatrix}, \tag{5.6}$$

with the unit vector

$$\boldsymbol{n} = \frac{\boldsymbol{x}_2 - \boldsymbol{x}_1}{\|\boldsymbol{x}_2 - \boldsymbol{x}_1\|}. \tag{5.7}$$

Note that the matrices $\boldsymbol{A}_{1,2}$ and $\boldsymbol{B}_{1,2}$ satisfy

$$\begin{cases} {\boldsymbol{A}_{1,2}}^2 = 2\boldsymbol{A}_{1,2}, \\ {\boldsymbol{B}_{1,2}}^2 = 2\boldsymbol{B}_{1,2}, \\ \boldsymbol{A}_{1,2}\boldsymbol{B}_{1,2} = \boldsymbol{B}_{1,2}\boldsymbol{A}_{1,2} = \boldsymbol{0}. \end{cases} \tag{5.8}$$

Therefore, for two given positive coefficients $\tilde{\alpha}$ and $\tilde{\beta}$ which may depend on $r_{1,2}$, a pair of $\widetilde{\boldsymbol{H}}_{1,2}$ and $\widetilde{\boldsymbol{\Xi}}_h$ satisfying Eq (5.4) can be given as

$$\begin{cases} \widetilde{\boldsymbol{H}}_{1,2} = \tilde{\alpha}\boldsymbol{A}_{1,2} + \tilde{\beta}\boldsymbol{B}_{1,2}, \\ \widetilde{\boldsymbol{\Xi}}_h = \left\{ \sqrt{k_B T \tilde{\alpha}}\boldsymbol{A}_{1,2} + \sqrt{k_B T \tilde{\beta}}\boldsymbol{B}_{1,2} \right\} \begin{bmatrix} \boldsymbol{\xi}_1^G \\ \boldsymbol{\xi}_2^G \end{bmatrix}, \end{cases} \tag{5.9}$$

with the normalized Gaussian noise satisfying

$$\langle \begin{bmatrix} \boldsymbol{\xi}_1^G \\ \boldsymbol{\xi}_2^G \end{bmatrix} \otimes \begin{bmatrix} \boldsymbol{\xi}_1^G \\ \boldsymbol{\xi}_2^G \end{bmatrix} \rangle = \boldsymbol{I}_{3\times 2}. \tag{5.10}$$

From (5.8), the eigenspaces of $\boldsymbol{H}_{1,2}$ are composed of those of $\boldsymbol{A}_{1,2}$ and $\boldsymbol{B}_{1,2}$. The null space is spanned by the parallel displacements $\begin{bmatrix} \boldsymbol{d} \\ \boldsymbol{d} \end{bmatrix}, {}^\forall \boldsymbol{d}$. The eigenspace of $\boldsymbol{A}_{1,2}$ is spanned by the vector $\begin{bmatrix} \boldsymbol{n} \\ -\boldsymbol{n} \end{bmatrix}$, which corresponds to oscillation along $\boldsymbol{n}$. The eigenspace of $\boldsymbol{B}_{1,2}$ is spanned by the vectors $\begin{bmatrix} \boldsymbol{m} \\ -\boldsymbol{m} \end{bmatrix}$, $\boldsymbol{m} \perp \boldsymbol{n}$, which corresponds to the rotation about the center $(\boldsymbol{x}_1 + \boldsymbol{x}_2)/2$ (Figure 3). Because the nonzero eigenvalues of $\boldsymbol{A}_{1,2}$ and $\boldsymbol{B}_{1,2}$ are equal to 2, the nonzero eigenvalues of $\boldsymbol{H}_{1,2}$ are $2\ddot{\varphi}\big(r_{1,2}\big)$ and $2\,\dot{\varphi}\big(r_{1,2}\big)/r_{1,2}$. In a case of a bond between a pair of adjacent particles $(i, i+1)$,

$$\varphi(r) = c_B(r - r_0)^2, \tag{5.11}$$

The coefficient of $\boldsymbol{B}_{1,2}$ in (5.2) is negative for $r < r_0$. In this paper, we replace $(r - r_0)/r$ with a nonnegative constant b_B as follows:

$$\begin{cases} \tilde{\alpha}_B(r) = 2c_B, \\ \tilde{\beta}_B(r) = 2c_B \max((r - r_0)/r\,, b_B)\,. \end{cases} \tag{5.12}$$

Hereafter, we refer to the parameter b_B as SimHec(b_B) or SimHec-RC(b_B). In the following, a nearly optimal parameter is chosen for each approach.

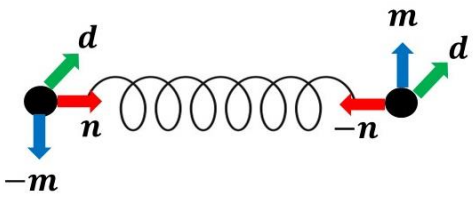

Figure 3. Eigenvectors of $\boldsymbol{A}_{1,2}$ and $\boldsymbol{B}_{1,2}$. Parallel displacements (green arrows) span the null space. Oscillations in the direction of $\boldsymbol{n}$ (red arrows) span the eigenspace of $\boldsymbol{A}_{1,2}$. Rotations represented by $\boldsymbol{m}$ orthogonal to $\boldsymbol{n}$ (blue arrows) span the eigenspace of $\boldsymbol{B}_{1,2}$.

6. Validation on entropic elasticity problems

To test the efficiency and the accuracy of SimHEC-RC, we simulated the entropic elasticity of a freely jointed chain consisting of n particles ($n = 100$) with potential

$$U(\boldsymbol{X}) = \sum_{i=1}^{n-1} c_B\,(\|\boldsymbol{x}_i - \boldsymbol{x}_{i+1}\| - r_{+1})^2. \tag{6.1}$$

We adopted the parameters of the coarse-grained protein model CafeMol [11], which only considers the position of a single bead at the C_α position for each residue in the polymer chain. A friction coefficient of $\gamma_i = 168.7$ CafeMolTime $\cdot$ kcal/(mol Å^2) was used. This value was derived by assuming the viscosity of water: 8×10^{-4} kg/(m s), and the radius r_{+1} of the amino acids: 3.82 Å. The time unit in CafeMol (CafeMolTime) is approximately 49 fs. A value of $c_B = 110.4$ kcal/(mol Å^2) was used for all bonds; the limitation on the time step size of the EM scheme due to the bond stiffness was estimated as $h \leq \gamma/2c_B$ =0.76 CafeMolTime.

First, we examined the accuracy for a transient problem where one end of the straight chain was released (Figure 4A). In this numerical test, one end $\boldsymbol{x}_1$ was always fixed, and the other end $\boldsymbol{x}_n$ was released at $t = 0$. The trajectories of the coordinates of $\boldsymbol{x}_n$ in the stretched direction were analyzed using the numerical results of 64 samples. We observed that the time transients of edge coordinate averages with 64 samples and 256 samples were almost identical in SimHec-RC with $h = 100$. Therefore, we performed the comparisons with 64 samples. The EM scheme with $h = 0.5$ shows a severe discrepancy (Figure 4B). The SimHec and SimHec-RC schemes were compared with the EM scheme with $h = 0.125$ (Figure 4C). The accuracy of SimHec deteriorated as the time step h increased. The shortening speed decreased with excessive friction as shown in (4.4) and (4.9). However, SimHec-RC retained its accuracy by correcting the random force.

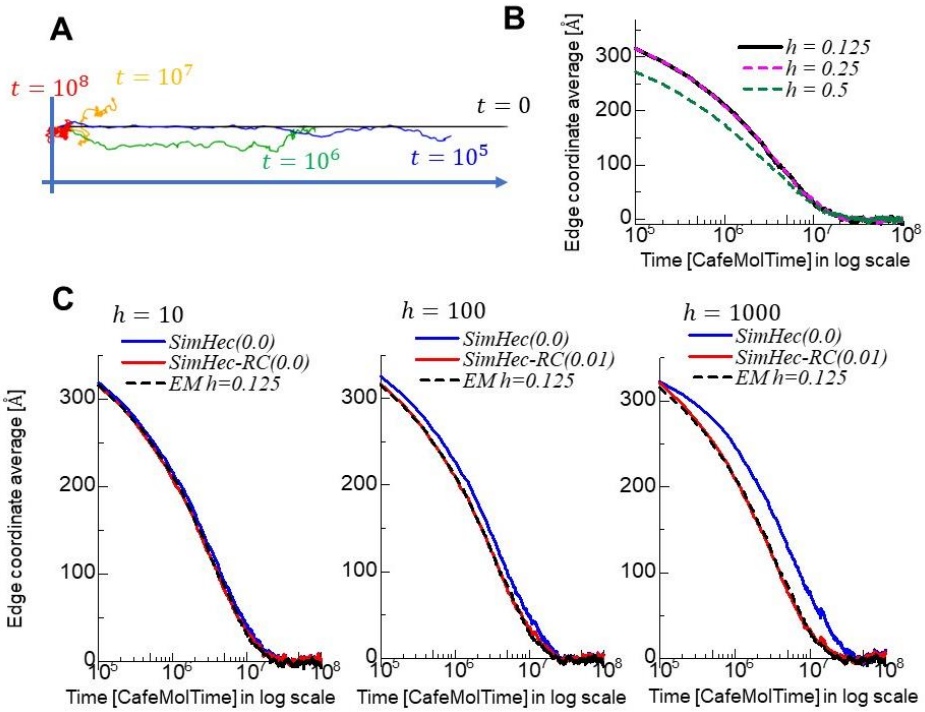

Figure 4. Time transients of the free end of the freely jointed chain. The averages of the free end coordinates of 64 samples are plotted. (A) Examples of the chain shape at various times. (B) The averaged coordinates of the free end calculated using EM schemes with $h =$ 0.25 (broken pink), 0.5 (broken green), and h =0.125 (black). (C) The coordinates of the free end as calculated using SimHec-RC (red) and SimHec (blue) schemes are compared with those using the EM scheme with $h =$ 0.125 (broken black) for $h =$ 10 (left panel), 100 (center panel), and 1000 (right panel). In SimHec-RC, b_B =0.0 was adopted for $h =$ 10, and b_B =0.01 was adopted for $h =$ 100 and 1000.

To confirm the accuracy for multiple steps as estimated in Figures 1 and 2, we examined the distributions of $(x, y) = (r\cos\theta, r\sin\theta)$, where r and θ are defined using the changes of $\boldsymbol{x}_{i+nd} - \boldsymbol{x}_i$ from t to $t + kh$ as follows:

$$\begin{cases} r = \|\boldsymbol{x}_{i+nd}(t+kh) - \boldsymbol{x}_i(t+kh)\|, \\ \cos\theta = \dfrac{\boldsymbol{x}_{i+nd}(t+kh) - \boldsymbol{x}_i(t+kh)}{\|\boldsymbol{x}_{i+nd}(t+kh) - \boldsymbol{x}_i(t+kh)\|} \cdot \dfrac{\boldsymbol{x}_{i+nd}(t) - \boldsymbol{x}_i(t)}{\|\boldsymbol{x}_{i+nd}(t) - \boldsymbol{x}_i(t)\|}, 0 \le \theta < \pi. \end{cases} \tag{6.2}$$

In Figure 5, we compared the distributions superposed for all pairs $(i, i + nd)$ with nd = 10 obtained using the EM scheme with $h = 0.125$ and SimHec-RC(0.01) with $h = 100$. There are substantial differences between the $k = 1$ and 10 cases of SimHec-RC and their counterparts, namely, the $k = 800$ and 8000 cases of the EM scheme. However, as expected from Figure 1A and B, the difference is diminished for $k = 100$.

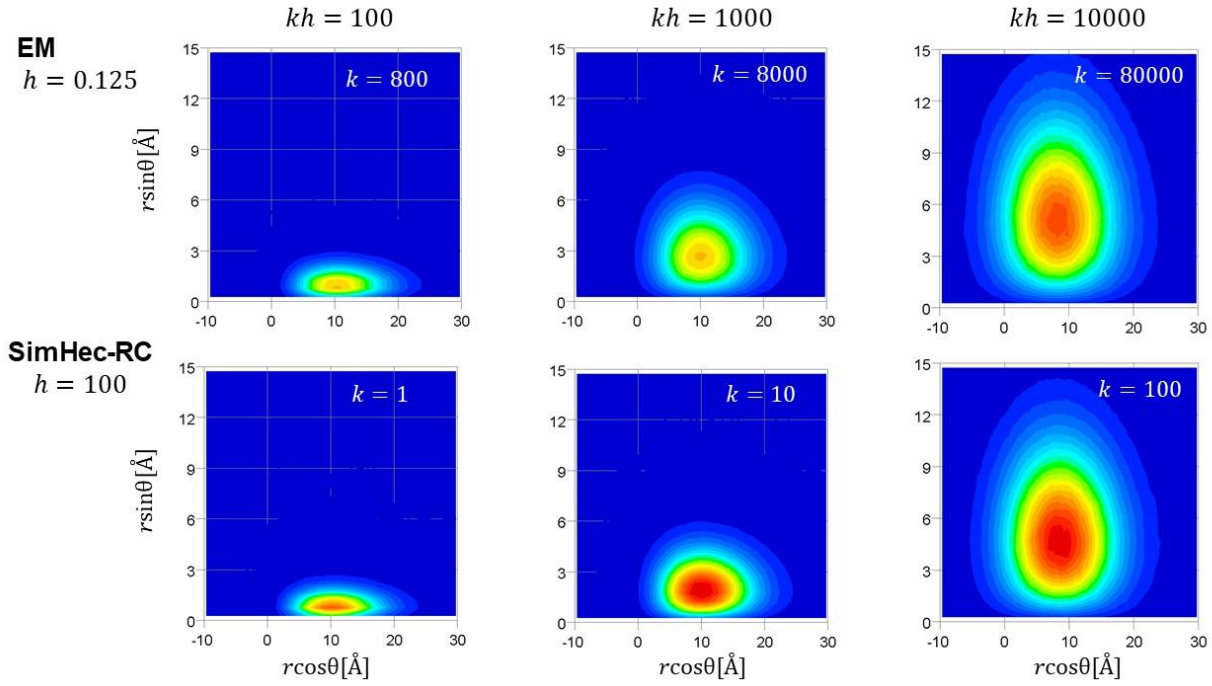

Figure 5. Distributions of changes of $\boldsymbol{x}_{i+10} - \boldsymbol{x}_i$ in the EM scheme (top) with h = 0.125 and SimHec-RC(0.01) (bottom) with h = 100 for kh = 100 (left), 1000 (center), and 10000 (right). Red indicates high concentration. The same color scale is used in each plot with the same time interval kh.

Note that hk = 10^4CafeMolTime ~ 0.5 ns is much smaller than the total time 10^7 CafeMolTime ~ 0.5 μs, which needed to be reduced to close to zero (Figure 4).

Next, we evaluated the force acting on the end $\boldsymbol{x}_n$ trapped by the spring at the position $\boldsymbol{p}_L = [nr_{+1}, 0,0]^T$ as

$$U(\boldsymbol{X}) = \sum_{i=1}^{n-1} c_B \left(\|\boldsymbol{x}_i - \boldsymbol{x}_{i+1}\| - r_{+1}\right)^2 + K_a\|\boldsymbol{x}_n - \boldsymbol{p}_L\|^2. \tag{6.3}$$

In the numerical experiment, we adopted K_a=0.1 kcal/(mol Å^2). In Figure 6A, the average values of the force

$$\langle f_1 \rangle = 2K_a\langle nr_{+1} - x_{n,1}\rangle \tag{6.4}$$

computed over the time range $[0,5 \times 10^9]$ CafeMolTime (~[0,0.25]s) for a single process are compared with the values calculated using the explicit scheme with $h = 0.25$. The EM scheme with $h = 0.5$ shows a severe discrepancy (Figure 6B) as in the case of shortening of the free end. In Figure 6C, we compared the average force as computed by SimHec and SimHec-RC, where c_B and K_a were increased by factors of one hundred and ten, respectively, with the theoretical values in the limit of $c_B \to \infty$:

$$\langle f_1 \rangle_\infty(L) = \frac{k_B T}{r_{+1}} F^{-1}\left(\frac{L}{nr_{+1}}\right), \tag{6.5}$$

where the Langevin function F is given by

$$F(x) = \frac{e^x + e^{-x}}{e^x - e^{-x}} - \frac{1}{x}. \tag{6.6}$$

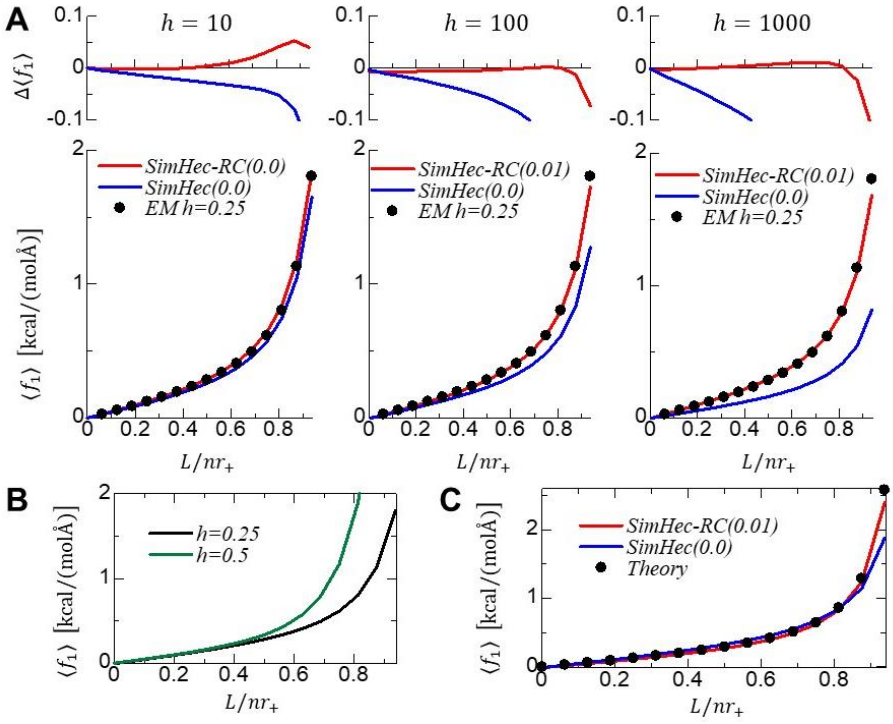

Figure 6. Entropic elasticity (A) Average values of the force $\langle f_1 \rangle$ acting on the end point (bottom), and the difference compared with the EM scheme with $h = 0.25$ (top). SimHec-RC (red) and SimHec (blue) are compared with EM with $h = 0.25$ (black points) for $h = 10$ (left panel), 100 (center panel), and 1000 (right panel). (B) EM with $h = 0.5$ (green) is compared with $h = 0.25$ (black). (C) SimHec-RC (red) and SimHec (blue) for the highly stiff bond parameter $c_B = 110.4 \times 100$ kcal/(mol Å^2) are compared with the theoretical limit where $c_B \to \infty$.

7. Computational efficiency

To assess the computational efficiency, we compared the overhead of SimHec-RC with that of EM. The overheads consist of the cost of computing the elemental Hessian matrices, the random force corrections, and the linear solutions. The computation times per time step for EM, SimHec, and SimHec-RC were 3.2, 16.9, and 22.6 μs, respectively, for a simulation of 100 particles using an Intel(R) Xeon(R) Gold 6448Y (4.1 GHz) single core processor. In the linear solution, the band structure with bandwidth 5 was exploited in the Cholesky factorization. The overhead of SimHec-RC was larger than that of SimHec owing to the random force correction. Comparing EM with $h = 0.25$ and SimHec-RC with $h = 100$ (resp. $h = 1000$), we found a speedup of a factor of 56.6 (resp. 566).

8. Conclusions and future work

In this study, we analyzed the relationship between the corrected Hessian matrix in a semi-implicit scheme and the random force concerning the accuracies for both the linear problem and the derived FP equations of the nonlinear problem. The proposed correction to the random force worked well for the simple MD problem involving bonded interactions. Although our approach is potentially extendable to more general potentials with bond angles and dihedral angles [3], the appropriate corrections to the Hessian matrix must be developed to avoid negative eigenvalues.

Author contributions

Toshiaki Hisada designed the project; Takumi Washio and Akihiro Fujii developed the semi-implicit Hessian correction scheme with random force correction, ran the simulations, analyzed the simulation data, and wrote the manuscript with input from Toshiaki Hisada; Takumi Washio developed the simulation code with input from Akihiro Fujii and Toshiaki Hisada. All authors contributed to the article and approved the submitted version.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work was supported by MEXT as "Program for Promoting Researches on the Supercomputer Fugaku" (Project ID: hp230216). We thank Edanz (https://jp.edanz.com/ac) for editing a draft of this manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

1. K. Bathe, A. Cimento, Some practical procedures for the solution of nonlinear finite element equations, *Comput. Method. Appl. M.*, **22** (1980), 59–85. https://doi.org/10.1016/0045-7825(80)90051-1
2. N. Schafer, D. Negrut, A quantitative assessment of the potential of implicit integration methods for molecular dynamics simulation, *J. Comput. Nonlin. Dyn.*, **5** (2010), 031012. https://doi.org/10.1115/1.4001392
3. T. Washio, R. Kanada, X. Cui, J. Okada, S. Sugiura, S. Takada, et al., Semi-implicit time integration with Hessian eigenvalue corrections for a larger time step in molecular dynamics simulations, *J. Chem. Theory Comput.*, **17** (2021), 5792–5804. https://doi.org/10.1021/acs.jctc.1c00398

4. P. Kloeden, E. Platen, *Numerical solution of stochastic differential equations*, Berlin: Springer, 1992. https://doi.org/10.1007/978-3-662-12616-5
5. N. Grønbech-jensen, S. Doniach, Long-time overdamped Langevin dynamics of molecular chains, *J. Comput. Chem.*, **15** (1994), 997–1012. https://doi.org/10.1002/jcc.540150908
6. C. Sweet, P. Petrone, V. Pande, J. Izaguirre, Normal mode partitioning of Langevin dynamics for biomolecules, *J. Chem. Phys.*, **128** (2008), 145101. https://doi.org/10.1063/1.2883966
7. C. Gardiner, *Handbook of stochastic methods*, Berlin: Springer, 1985. https://doi:10.1007/978-3-662-02377-8
8. H. Risken, *The Fokker Planck equation: methods of solution and applications*, Berlin: Springer, 1996. https://doi.org/10.1007/978-3-642-61544-3
9. W. Coffey, Y. Kalmykov, J. Waldron, *The Langevin equation: with applications to stochastic problems in physics, chemistry, and electrical engineering*, Singapore: World Scientific, 2004. https://doi.org/10.1142/9789812795090
10. K. Itô, *On stochastic differential equations*, New York: American Mathematical Society, 1951. https://doi.org/10.1090/memo/0004
11. CafeMolV3.2 CafeMol 3.2 manual, 2021. Available from: https://www.cafemol.org/doc/.