INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 1, Number 1, Pages 1–20

AN IMPROVED WPE METHOD FOR SOLVING DISCONTINUOUS FOKKER-PLANCK EQUATIONS

HONGYUN WANG

(Communicated by Steve Hou)

Abstract. In mathematical studies of molecular motors, the stochastic motor motion is modeled using the Langevin equation. If we consider an ensemble of motors, the probability density is governed by the corresponding Fokker-Planck equation. Average quantities, such as, average velocity, effective diffusion and randomness parameter, can be calculated from the probability density. The WPE method was previously developed to solve Fokker-Planck equations (H. Wang, C. Peskin and T. Elston, J. Theo. Biol., Vol. 221, 491-511, 2003). The WPE method has the advantage of preserving detailed balance, which ensures that the numerical method still works even when the potential is discontinuous. Unfortunately, the accuracy of the WPE method drops to first order when the potential is discontinuous. Here we propose an improved version of the WPE method. The improved WPE method a) maintains the second order accuracy even when the potential is discontinuous, b) has got rid of a numerical singularity in the WPE method, and c) is as simple and easy to implement as the WPE method. Numerical examples are shown to demonstrate the robust performance of the improved WPE method.

Key Words. Fokker-Planck equation, detailed balance, numerical solutions.

1. Introduction

Molecular motors are small, and, as a result, the motor operation is dominated by high viscous friction and large thermal fluctuations from the surrounding fluid environment [1]. In general, a molecular motor has many internal and external degrees of freedom. One of these degrees of freedom is associated with the motor's unidirectional motion, the main biological function of the motor. For example, a kinesin dimer walks along a microtubule toward the positive end [5, 6]. There are many levels of models for molecular motors, from simple kinetic models with a few states to all atom

Received by the editors February 20, 2006 and, in revised form, ??, 2006.
2000 *Mathematics Subject Classification.* 65Mxx, 82Cxx.

molecular dynamics. In a modeling approach of intermediate level, the unidirectional motion is followed explicitly and the effects of other degrees of freedom are modeled in the mean field potential affecting the unidirectional motion [7, 8, 9].

To introduce this modeling approach of intermediate level, we consider the one dimensional motion of a small object in water. The motion of the object is governed by the Newton's second law:

$$m\frac{dv}{dt} = -\zeta v - \phi'(x) + \sqrt{2k_BT\zeta}\,\frac{dW(t)}{dt} \tag{1}$$

where x is the coordinate along the dimension of motion, m is the mass and $v = \frac{dx}{dt}$ the velocity of the object, ζ is the drag coefficient, $\phi(x)$ a potential affecting the motion of object, and $W(t)$ is the Weiner process. The object is affected by a) the drag force $-\zeta v$, which is always opposing the motion, b) the force derived from the potential, and c) the Brownian force. Both the drag force and the Brownian force are caused by the bombardments of surrounding water molecules. The amplitude of Brownian force is related to the drag coefficient as $\sqrt{2k_BT\zeta}$, which is a result of the fluctuation-dissipation theorem [16, 17, 18]. Here k_B is the Boltzmann constant and T the absolute temperature [13].

In (1), there is a very short time scale associated with the object forgetting about its instantaneous velocity. It is called the time scale of inertia [11]. For both theoretical analysis and numerical solutions, it is more convenient to get rid of this short time scale and make the system non-stiff. We start by rewriting (1) as

$$\frac{dv}{dt} = -\frac{1}{t_0}\left[v - \left(-D\frac{\phi'(x)}{k_BT} + \sqrt{2D}\frac{dW(t)}{dt}\right)\right] \tag{2}$$

where $D = \frac{k_BT}{\zeta}$ is the diffusion coefficient, and $t_0 = \frac{m}{\zeta}$ has the dimension of time. In (2), in the absence of potential $\phi(x)$, the autocorrelation of the instantaneous velocity satisfies

$$\langle v(s)v(s+t)\rangle = \langle v^2(s)\rangle \exp\left(\frac{-t}{t_0}\right) \tag{3}$$

where $\langle\cdot\rangle$ denotes the average. It is clear that the object forgets about its current velocity after a small multiples of t_0. That is why t_0 is called the time scale of inertia. In the simple case where the object is a spherical bead of radius σ, the mass and the drag coefficient are respectively [1]

$$m = \frac{4}{3}\pi\rho\sigma^3 \quad , \qquad \zeta = 6\pi\eta\sigma \tag{4}$$

where ρ is the density and η the viscosity of water. The time scale $\frac{m}{\zeta}$ is proportional to the square of radius: $t_0 = \frac{m}{\zeta} = O(\sigma^2)$. Consequently, for

small objects, the time scale $t_0 = \frac{m}{\zeta}$ is extremely small. For a bead of $1\mu m$ in diameter, the time scale of inertia is $t_0 = 56 \times 10^{-9} s = 56ns$ [10].

When t_0 is very small, (2) is well approximated by

$$v = \left[-D\frac{\phi'(x)}{k_B T} + \sqrt{2D}\frac{dW(t)}{dt} \right] \tag{5}$$

The reduction from (2) to (5) in the limit of small t_0 is called the Einstein-Smoluchowski limit [16, 14]. This reduction can be illustrated intuitively by considering a simple model equation: $y' = -\frac{1}{t_0}(y - f(t))$. The exact solution of the model equation is given by

$$\begin{aligned} y(t) &= f(t) + \exp\left(\frac{-t}{t_0}\right)(y(0) - f(t)) \\ &\quad + \frac{1}{t_0}\int_0^t \exp\left(\frac{-(t-s)}{t_0}\right)(f(s) - f(t))\, ds \end{aligned} \tag{6}$$

When t_0 is small and $t >> t_0$, the exact solution satisfies approximately $y(t) = f(t)$, which is comparable to (5). Writing (5) as a differential equation for x, we obtain

$$\frac{dx}{dt} = -D\frac{\phi'(x)}{k_B T} + \sqrt{2D}\frac{dW(t)}{dt} \tag{7}$$

This is the Langevin equation, governing the stochastic motion of a small object affected by a static potential $\phi(x)$ [14].

In molecular motors, however, the potential is not static. In molecular motors, the current potential is determined by the current chemical occupancy state of the motor. The general mathematical framework used in modeling molecular motors is a collection of Langevin equations. Each Langevin equation in the collection corresponds to a chemical state and has the form of (7) with a periodic potential $\phi_S(x)$ [7, 9, 4]:

$$\frac{dx}{dt} = -D\frac{\phi'_S(x)}{k_B T} + \sqrt{2D}\frac{dW(t)}{dt}, \tag{8}$$

where S represents the current chemical state of the motor. At any given time, the motor is governed by one Langevin equation in the collection. When the motor switches to another chemical state, the governing Langevin equation changes accordingly. The period of these potentials is usually a small multiples of the motor step. For a kinesin dimer walking on microtubule, the step size is about 8 nm [6]. The chemical reaction of the motor system is governed by a discrete space Markov process (a jump process).

In experiments, only average quantities can be measured reliably. All average quantities (including the average velocity and the effective diffusion) can be calculated efficiently by following the probability density of the motor. Let us consider an ensemble of motors, each evolving in time independently and stochastically according to Langevin equation (8) coupled

with the discrete Markov process describing the reaction. Let $\rho_S(x,t)$ be the probability density that the motor is at position x and in chemical state S at time t. $\rho_S(x,t)$ is governed by the Fokker-Planck equation corresponding to Langevin equation (8) [14]:

$$\frac{\partial \rho_S}{\partial t} = D\frac{\partial}{\partial x}\left[\frac{\phi_S'(x)}{k_BT}\rho_S + \frac{\partial \rho_S}{\partial x}\right] + \sum_{j=1}^{N} k_{j\to S}(x)\rho_j, \qquad S = 1, 2, ..., N \tag{9}$$

where, N is the total number of chemical states of the motor system. For $j \neq S$, $k_{j\to S}(x)$ is the chemical transition rate from state j to state S. $k_{S\to S}(x)$ is the total rate of jumping out of state S and is given by

$$k_{S\to S}(x) = -\sum_{j\neq S} k_{S\to j}(x) \tag{10}$$

Thus, the transition matrix $\{k_{S\to j}(x)\}$ always satisfies $\sum_{j=1}^{N} k_{S\to j}(x) = 0$.

The motor operation is governed by Fokker-Planck equation (9) with a set of N potentials, one for each chemical state. A simpler way to model molecular motors is to characterize the motor operation using a single potential obtained by averaging $\phi_S'(x)$ over all chemical states weighted by the steady state probability density functions of these states [10]. Let $\psi'(x)$ be the weighted average of $\phi_S'(x)$ over all chemical states defined as:

$$\psi'(x) = \frac{1}{\rho(x)}\sum_{S=1}^{N} \rho_S(x)\phi_S'(x) \quad , \qquad \rho(x) = \sum_{S=1}^{N} \rho_S(x) \tag{11}$$

where $\rho(x)$ is the probability density that the motor is at position x at time t, regardless of the chemical state, and $\psi(x)$ is called the motor potential profile. The biggest advantage of using motor potential profile $\psi(x)$ to characterize the motor operation is that $\psi(x)$ can be extracted from time series of motor positions measured in single molecule experiments [11]. Thus, essentially, $\psi(x)$ is a measurable entity (it is a function, instead of a number). The motor potential profile $\psi(x)$ can be viewed as the motor's mean field free energy landscape. Let L be the period of $\phi_S(x)$. We immediately see that $\psi'(x)$ is also periodic with period L. As a result, the motor potential profile $\psi(x)$ is a tilted periodic potential: $\psi(x+L) = \psi(x) - \Delta\psi$, where $\Delta\psi > 0$ is the energy made available from the chemical reaction to driving the motor forward in one period. The mechanical motion of the motor can be modeled approximately using Langevin equation (7) with potential $\psi(x)$. The corresponding Fokker-Planck equation is [14]:

$$\frac{\partial \rho}{\partial t} = D\frac{\partial}{\partial x}\left[\frac{\psi'(x)}{k_BT}\rho + \frac{\partial \rho}{\partial x}\right] \tag{12}$$

This approach of using a single tilted periodic potential to model the motor operation has been used in studies of motors [19, 20].

In [21], a robust numerical method was proposed for solving Fokker-Planck equations (9) and (12). Below, this numerical method will be referred to as the WPE method. When the potential is smooth, the WPE method is convergent with second order accuracy [21]. The advantage of the WPE method is that it works fairly well even if the potential is discontinuous. When the potential is discontinuous, numerical simulations showed that the WPE method is convergent with first order accuracy, which has been proved mathematically recently [12]. In spite of the robust performance of the WPE method, it has two aspects on which we like to improve: a) the accuracy of the WPE method drops from second order to first order when the potential is discontinuous, and b) the jump rates of the WPE method have a numerical singularity that requires different numerical formulas for different numerical values.

In a convergence study of the WPE method [12], we analyzed the effect of the discontinuity on the truncation error of the WPE method. Based on the analysis, we proposed a way to improve the WPE method for the case where the discontinuity of potential is at a half grid point. Since the focus of [12] was to prove the convergence (mainly the stability of the WPE method), we did not pursue it very far on how to improve the WPE method. It turns out that the rudimentary idea for improving the WPE method in [12] works only for the case where the discontinuity of potential is at a half grid point. In this paper, we develop a new strategy to improve the WPE method to preserve detailed balance and to achieve second order accuracy even if the discontinuity of potential is not at a half grid point. This is achieved by a new way of discretizing the probability flux, in which, two new free energy concepts are introduced: the free energy of a cell and the free energy at a border between two cells. The free energy of cell is introduced to preserve detailed balance while the free energy of border is formed to exploit the continuity property of the probability flux and other functions to achieve second order accuracy. Here we like to point out the importance of achieving second order accuracy for the case where the discontinuity of potential is not at a half grid point. Although in a 1-D problem, we can always shift the numerical grid to accommodate one discontinuity, it is not so easy to accommodate multiple discontinuities. For 2-D problems it is virtually impossible to make discontinuities reside on half grid points. While preserving detailed balance and achieving second order accuracy, the improved WPE method is also as simple and easy to implement as the WPE method. As a matter of fact, because we have got rid of the numerical singularity, the evaluation of jump rates no longer requires different numerical formulas for different situations. Consequently it is easier to implement the improved WPE method and more straightforward to port it to parallel environment than the WPE method.

The rest of the paper is organized as follows. In section 2, we describe the construction and properties of the WPE method. The most important property of the WPE method is detailed balance, which ensures that the method works even if the potential is discontinuous. In section 3, we view the Fokker-Planck equations (9) and (12) in a slightly different way, and propose an improved version of WPE method. The improved WPE method is always second order accurate even if the potential is discontinuous. More specifically, when the discontinuity is at a half grid point, the improved WPE method is second order in L^1 norm, L^2 norm and L^∞ norm. When the discontinuity is not at a half grid point, the improved WPE method is second order in L^1 norm and all average quantities obtained by the improved WPE method (including average velocity and effective diffusion) have second order accuracy. The improved WPE method preserves both the second order accuracy and detailed balance. In section 4, we analyze the truncation error of the improved WPE method. We will show that although the truncation error of the improved WPE method with respect to the exact solution is $O(1)$, the truncation error of the improved WPE method with respect to a perturbed solution is second order. The difference between the perturbed solution and the exact solution is proportional to the square of the spatial step. As a result, the spatial accuracy of the improved WPE method is second order. For the time integration, we use a two stage second order diagonally implicit Runge Kutta (DIRK) method [15]. In comparison with the frequently used Crank-Nicolson method (trapezoidal method) [24], the DIRK method is L-stable while the Crank-Nicolson method is only A-stable. An A-stable method can prevent high wave number oscillations from growing but it is not good at suppressing high wave number oscillations once they appear. An L-stable method is capable of suppressing high wave number oscillations. Thus, using an L-stable DIRK method for time integration here is consistent with the diffusive property of the underlying physical process. In section 5, we carry out numerical simulations to compare the performance of the improved WPE method and the WPE method.

2. The WPE method

In this section, we review the WPE method developed in [21]. For simplicity, we use (12) as a model equation in the description of numerical methods. The extension of numerical methods to Fokker-Planck equations with reactions (such as (9)) is straightforward.

(12) is the Fokker-Planck equation for the underlying Langevin equation (7), which is a continuous time continuous space Markov process. The philosophy of the WPE method is that in the spatial discretization, the Langevin equation is approximated by a continuous time discrete space Markov process (a jump process), and consequently, the corresponding Fokker-Planck equation is approximated by the master equation of the jump

process. More specifically, we start by dividing the period $[0, L]$ into M subintervals of size $h = L/M$. The numerical grid is formed as

$$h = \frac{L}{M}, \quad x_j = \frac{h}{2} + jh, \quad x_{j+1/2} = \frac{h}{2} + (j + \frac{1}{2})h$$

Each subinterval $(x_{j-1/2}, x_{j+1/2})$ is called a cell and is represented by its center x_j. Each cell center x_j is also called a site. Two adjacent cells x_j and x_{j+1} are separated by a border $x_{j+1/2}$ as shown in Figure 1. Since each cell is represented by a site, the motor is restricted to the set of sites $\{x_j\}$. In a single jump, it is only allowed to jump to an adjacent site.

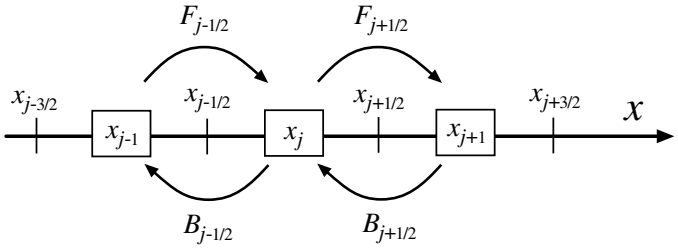

FIGURE 1. Spatial discretization of (12). In the jump process, the motor is only allowed to jump from one cell to an adjacent cell at a time.

Let $p_j(t)$ be the probability that the motor is at site x_j at time t in the jump process. Since the site x_j represents the subinterval $(x_{j-1/2}, x_{j+1/2})$, $p_j(t)$ can be viewed as

$$p_j(t) \approx \int_{x_{j-1/2}}^{x_{j+1/2}} \rho(x,t)dx \approx h \cdot \rho(x_j, t) \tag{13}$$

Let $F_{j+1/2}$ be the rate of jumping from x_j to x_{j+1} (Forward jump), $B_{j+1/2}$ the rate of jumping from x_{j+1} to x_j (Backward jump). The numerical probability flux through $x_{j+1/2}$ is

$$J_{j+1/2} = F_{j+1/2}\, p_j - B_{j+1/2}\, p_{j+1} \tag{14}$$

The time evolution of $p_j(t)$ is governed by the conservation of probability:

$$\begin{aligned} \frac{dp_j}{dt} &= J_{j-1/2} - J_{j+1/2} \\ &= (F_{j-1/2}\, p_{j-1} - B_{j-1/2}\, p_j) - (F_{j+1/2}\, p_j - B_{j+1/2}\, p_{j+1}) \end{aligned} \tag{15}$$

In the WPE method [21], the jump rates are calculated from local approximate steady state solutions. Below, for simplicity, we assume $k_B T = 1$, which is equivalent to that potential $\psi(x)$ has been normalized by $k_B T$. For simplicity, we also assume $D = 1$. In the interval $[x_{j-1/2}, x_{j+3/2}]$, which

spans cells x_j and x_{j+1}, we use a linear function to approximate potential $\psi(x)$:

$$\psi(x) \approx \psi_{j+1/2}(x) \equiv C + \frac{x}{h}\delta\psi_{j+1/2} \tag{16}$$

where

$$\delta\psi_{j+1/2} = \psi(x_{j+1}) - \psi(x_j) \tag{17}$$

Let $\rho_{j+1/2}(x)$ denote the local steady state solution of (12) in $[x_{j-1/2}, x_{j+3/2}]$ with linear potential (16) and subject to the constraints:

$$\begin{aligned} \int_{x_{j-1/2}}^{x_{j+1/2}} \rho_{j+1/2}(x)dx &= p_j \\ \int_{x_{j+1/2}}^{x_{j+3/2}} \rho_{j+1/2}(x)dx &= p_{j+1} \end{aligned} \tag{18}$$

Solving for $\rho_{j+1/2}(x)$, we obtain [21]

$$\begin{aligned} \rho_{j+1/2}(x) \;=\; & \frac{\delta\psi_{j+1/2}\,\exp(\delta\psi_{j+1/2})(p_j - p_{j+1})}{(\exp(\delta\psi_{j+1/2}) - 1)^2}\,\exp\left(-\delta\psi_{j+1/2}\frac{(x - x_{j+1/2})}{h}\right) \\ & -\frac{p_j - \exp(\delta\psi_{j+1/2})p_{j+1}}{(\exp(\delta\psi_{j+1/2}) - 1)} \end{aligned} \tag{19}$$

The probability flux of $\rho_{j+1/2}(x)$ is given by

$$J = \frac{1}{h^2}\cdot\frac{\delta\psi_{j+1/2}}{\exp(\delta\psi_{j+1/2}) - 1}\left(p_j - \exp(\delta\psi_{j+1/2})\,p_{j+1}\right) \tag{20}$$

Comparing the theoretical flux (20) with the numerical flux (14), we obtain immediately

$$\begin{aligned} F_{j+1/2} &= \frac{1}{h^2}\cdot\frac{\delta\psi_{j+1/2}}{\exp(\delta\psi_{j+1/2}) - 1} \\ B_{j+1/2} &= \frac{1}{h^2}\cdot\frac{\delta\psi_{j+1/2}\,\exp(\delta\psi_{j+1/2})}{\exp(\delta\psi_{j+1/2}) - 1} \end{aligned} \tag{21}$$

The jump rates given in (21) for the WPE method have a numerical singularity. Specifically, when $\delta\psi_{j+1/2} = 0$, both the numerator and the denominator in (21) are zero so a straightforward evaluation using (21) will yield $0/0$. When $\delta\psi_{j+1/2}$ is very close to zero, the denominator is the difference between two terms that are nearly equal to each other. In this case, a straightforward evaluation using (21) will lose numerical accuracy due to numerical cancellation. Thus, when $\delta\psi_{j+1/2}$ is very close to zero (say $|\delta\psi_{j+1/2}| < 10^{-3}$), we Taylor expand (21) and use the numerical formula

below to evaluate the jump rates.

$$F_{j+1/2} = \frac{1}{h^2} \cdot \frac{1}{1 + \frac{1}{2!}\delta\psi_{j+1/2} + \frac{1}{3!}(\delta\psi_{j+1/2})^2 + \frac{1}{4!}(\delta\psi_{j+1/2})^3}$$

$$B_{j+1/2} = \frac{1}{h^2} \cdot \frac{\exp(\delta\psi_{j+1/2})}{1 + \frac{1}{2!}\delta\psi_{j+1/2} + \frac{1}{3!}(\delta\psi_{j+1/2})^2 + \frac{1}{4!}(\delta\psi_{j+1/2})^3} \tag{22}$$

We now show that the jump rates (21) preserve detailed balance. Detailed balance is a thermodynamic property of the underlying physical system: if the system is brought to an equilibrium, then the probability distribution is given by the Boltzmann distribution and the probability flux vanishes everywhere [13]. Substituting the Boltzmann distribution $p_j^{(e)} \propto \exp(-\psi(x_j))$ into the numerical flux (14), we have

$$\begin{aligned} J_{j+1/2}^{(e)} &= F_{j+1/2}\, p_j^{(e)} - B_{j+1/2}\, p_{j+1}^{(e)} \\ &= B_{j+1/2}\, p_j^{(e)} \left[\frac{F_{j+1/2}}{B_{j+1/2}} - \frac{p_{j+1}^{(e)}}{p_j^{(e)}} \right] \\ &= B_{j+1/2}\, p_j^{(e)} \left[\frac{F_{j+1/2}}{B_{j+1/2}} - \exp(-\delta\psi_{j+1/2}) \right] \end{aligned} \tag{23}$$

The flux vanishes everywhere if and only if the jump rates satisfy

$$\frac{F_{j+1/2}}{B_{j+1/2}} = \exp(-\delta\psi_{j+1/2}) \tag{24}$$

Thus, (24) is equivalent to detailed balance. It is straightforward to verify that the jump rates given in (21) satisfy detailed balance (24).

3. The improved WPE method

In this section, we propose an improved version of WPE method. The improved WPE method is always second order accurate even if the potential is discontinuous. The improved WPE method preserves both the second order accuracy and detailed balance. This is achieved by a new way of approximating the probability flux. In approximating the probability flux, two new free energy concepts are introduced: the free energy of a cell and the free energy at a border between two cells. The free energy of cell is used to preserve detailed balance and the free energy of border is used in the finite difference to achieve high order accuracy.

3.1. Properties of exact solution at a discontinuity. We start by writing the probability flux of Fokker-Planck equation (12) in a slightly different form. As we will see the numerical approximation of the probability flux based on this form preserves detailed balance automatically. Again, for simplicity, we assume $k_B T = 1$, which is equivalent to that potential $\psi(x)$ has

been normalized by k_BT. For simplicity, we also assume $D = 1$. Writing (12) in the form of conservation of probability, we have

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x} J(x,t) \tag{25}$$

where the probability flux is

$$\begin{aligned} J(x,t) &= -\left[\psi'(x)\rho + \frac{\partial \rho}{\partial x}\right] \\ &= -\frac{1}{\exp(\psi(x))} \frac{\partial \left[\exp(\psi(x))\rho(x,t)\right]}{\partial x} \end{aligned} \tag{26}$$

Suppose potential $\psi(x)$ has a discontinuity at $x = d$. Let us examine the conditions that the exact solution $\rho(x,t)$ must satisfy at $x = d$. Away from $x = d$, potential $\psi(x)$ is smooth and the solution $\rho(x,t)$ satisfies the differential equation. At $x = d$, potential $\psi(x)$ is discontinuous. If reflecting boundary conditions are set properly to block flux, the system will reach an equilibrium and the equilibrium probability density is given by the Boltzmann distribution:

$$\rho^{(e)}(x) \propto \exp(-\psi(x)) \tag{27}$$

Since $\psi(x)$ is discontinuous at $x = d$, $\rho^{(e)}(x)$ is also discontinuous. Thus, in general, we should expect $\rho(x,t)$ to be discontinuous at $x = d$. Even if we start with $\rho(x,0)$ continuous at $x = d$, as time evolves, $\rho(x,t)$ will become discontinuous at $x = d$. At $x = d$, both $\psi(x)$ and $\rho(x,t)$ are discontinuous and the differential equation is not valid in the classical sense. To derive the conditions for $\rho(x,t)$ at $x = d$, we first look at the conservation of probability at $x = d$. The probability flux into $x = d$ and the probability flux out of $x = d$ must be the same: $J(d^-,t) = J(d^+,t)$. That is,

$$\frac{1}{\exp(\psi(x))} \frac{\partial \left[\exp(\psi(x))\rho(x,t)\right]}{\partial x}\bigg|_{x=d^-} = \frac{1}{\exp(\psi(x))} \frac{\partial \left[\exp(\psi(x))\rho(x,t)\right]}{\partial x}\bigg|_{x=d^+} \tag{28}$$

(28) is one of the two conditions for $\rho(x,t)$ at $x = d$. Since (25) is a second order equation, $\rho(x,t)$ needs two conditions at $x = d$. The second condition follows intuitively from that the probability flux at $x = d$ is finite and continuous for $t > 0$. Intuitively, the probability flux being finite at $x = d$ implies that $\exp(\psi(x))\rho(x)$ is continuous at $x = d$:

$$\exp(\psi(x))\rho(x,t)|_{x=d^-} = \exp(\psi(x))\rho(x,t)|_{x=d^+} \tag{29}$$

(29) is the second condition for $\rho(x,t)$ at $x = d$.

Strictly speaking, the probability flux at $x = d$ is not defined. The probability flux $J(x,t)$ is defined only for $x > d$ and $x < d$. The continuity of $\exp(\psi(x))\rho(x,t)$ at $x = d$ does not follow directly from $J(x)$ being finite for $x > d$ and $x < d$. The discontinuous transition of $\psi(x)$ at $x = d$ is

a mathematical abstraction. As a result, the solution has to be interpreted as the limit of solutions corresponding to a sequence of smooth potentials converging to the discontinuous potential. In the sequence of potentials, the discontinuity at $x = d$ is replaced by smooth transitions over smaller and smaller transition regions. To derive (29), we view the discontinuous transition of potential $\psi(x)$ at $x = d$ as a smooth transition $\psi_\varepsilon(x)$ over $[d-\varepsilon, d+\varepsilon]$ with ε converging to zero. In the transition region $[d-\varepsilon, d+\varepsilon]$, the potential $\psi_\varepsilon(x)$ is smooth and satisfies

$$\begin{aligned} &\lim_{\varepsilon\to 0} \psi_\varepsilon(d-\varepsilon) = \psi(d^-), \\ &\lim_{\varepsilon\to 0} \psi_\varepsilon(d+\varepsilon) = \psi(d^+) \end{aligned} \tag{30}$$

The exact solution $\rho_\varepsilon(x,t)$ corresponding to $\psi_\varepsilon(x)$ satisfies

$$\begin{aligned} &\lim_{\varepsilon\to 0} \rho_\varepsilon(d-\varepsilon,t) = \rho(d^-,t), \\ &\lim_{\varepsilon\to 0} \rho_\varepsilon(d+\varepsilon,t) = \rho(d^+,t) \end{aligned} \tag{31}$$

It is important to notice that since the potential is smooth the probability flux is well defined everywhere. In the transition region $[d-\varepsilon, d+\varepsilon]$, the probability flux is finite and the bound is independent of ε.

$$-\frac{1}{\exp(\psi_\varepsilon(x))}\frac{\partial\left[\exp(\psi_\varepsilon(x))\rho_\varepsilon(x,t)\right]}{\partial x} = J_\varepsilon(x,t) \tag{32}$$

Multiplying by $\exp(\psi_\varepsilon(x))$ and integrating from $d-\varepsilon$ to $d+\varepsilon$, we get

$$\exp(\psi_\varepsilon(x))\rho_\varepsilon(x,t)|_{x=d-\varepsilon}^{x=d+\varepsilon} = -\int_{d-\varepsilon}^{d+\varepsilon} \exp(\psi_\varepsilon(s))J_\varepsilon(s,t)ds \tag{33}$$

Taking the limit of both sides as $\varepsilon \to 0$ and using (30) and (31), we obtain

$$\exp(\psi(d^+))\rho(d^+,t) - \exp(\psi(d^-))\rho(d^-,t) = 0 \tag{34}$$

which leads immediately to condition (29).

Conditions (28) and (29) together will uniquely determine the correct solution of differential equation (25). Condition (28) is basically the conservation of probability. If a numerical method is based on conservation of probability, then the numerical solution will automatically satisfy condition (28). Condition (29), on the other hand, can be written as $\rho(d^\pm,t) \propto \exp(-\psi(d^\pm))$, which is the local Boltzmann distribution at $x = d$. Notice that the time scale of relaxing a region of size ε to equilibrium by diffusion is of the order ε^2 while the time scale of disturbing a region of size ε away from equilibrium by convection is of the order ε. Consequently, although the system in the big region is not in equilibrium or even not in a steady state, the system in each small sub-region of the size ε is always in quasi-equilibrium. That means local equilibrium is quickly established once it is disturbed. It follows that, at $x = d$ (a region of size zero), it is always in equilibrium. Thus, condition (29) is equivalent to detailed balance at

$x = d$, which ensures that at equilibrium the system has the Boltzmann distribution. If a numerical method does not preserve detailed balance, then the numerical solution may converge to a wrong solution that does not satisfy condition (29).

Condition (29) also implies that $\exp(\psi(x))\rho(x,t)$ is continuous at $x = d$ and everywhere. Thus, there are at least two advantages of expressing the probability flux in terms of the derivative of $\exp(\psi(x))\rho(x,t)$: a) the probability flux being zero is equivalent to the Boltzmann distribution $\rho(x) \propto \exp(-\psi(x))$ so detailed balance is automatically preserved, and b) applying finite difference on the continuous $\exp(\psi(x))\rho(x,t)$ (instead of on a discontinuous function) makes it possible to preserve the second order accuracy.

3.2. Spatial discretization of the improved WPE method. As in the WPE method, we form the numerical grid by dividing the period $[0, L]$ into M subintervals of size $h = L/M$. Each subinterval $(x_{j-1/2}, x_{j+1/2})$ is called a cell and is represented by its center x_j. Two adjacent cells x_j and x_{j+1} are separated by a border $x_{j+1/2}$, as illustrated in Figure 1.

Let $p_j(t)$ be a numerical approximation for the probability that the motor is in cell x_j at time t. Note that when we say cell x_j, we mean the subinterval $(x_{j-1/2}, x_{j+1/2})$. So $p_j(t)$ can be viewed as:

$$p_j(t) \approx \int_{x_{j-1/2}}^{x_{j+1/2}} \rho(x,t)dx \approx h \cdot \rho(x_j,t) \tag{35}$$

$p_j(t)$ is governed by the conservation of probability

$$\frac{dp_j(t)}{dt} = -(J_{j+1/2}(t) - J_{j-1/2}(t)) \tag{36}$$

where $J_{j+1/2}(t)$ is a numerical approximation for probability flux $J(x_{j+1/2}, t)$ given by:

$$J(x,t) = -\frac{1}{\exp(\psi(x))}\frac{\partial g(x)}{\partial x} \quad , \quad g(x) = \exp(\psi(x))\rho(x,t) \tag{37}$$

To preserve detailed balance, let us first define the free energy for each cell. Then in the spatial discretization, we will enforce that the equilibrium solution of the numerical method is related to the free energies of cells by the Boltzmann distribution. The equilibrium solution of differential equation (25) is $\rho^{(e)}(x) = \frac{1}{Z}\exp(-\psi(x))$ where Z is the normalizing factor in mathematics and is known as the partition function in physics. The corresponding numerical equilibrium solution should be

$$p_j^{(e)} = \frac{1}{Z}\int_{x_{j-1/2}}^{x_{j+1/2}} \exp(-\psi(s))ds \equiv \frac{h}{Z}\exp(-\psi_j^c) \tag{38}$$

where ψ_j^c is the free energy of cell x_j defined as

$$\psi_j^c = -\log\left(\frac{1}{h}\int_{x_{j-1/2}}^{x_{j+1/2}} \exp(-\psi(s))ds\right) \tag{39}$$

Here we use ψ_j^c to denote the free energy of cell x_j to distinguish it from ψ_j, which by convention is defined as $\psi_j = \psi(x_j)$. The exact value of the integral in definition (39) is usually difficult to evaluate. If the discontinuity is not in cell x_j, that is, $\psi(x)$ is smooth in $(x_{j-1/2}, x_{j+1/2})$, then we have $\psi_j^c = \psi_j + O(h^2)$. So in the numerical method, we simply use $\psi_j^c = \psi_j$ if cell x_j does not contain a discontinuity.

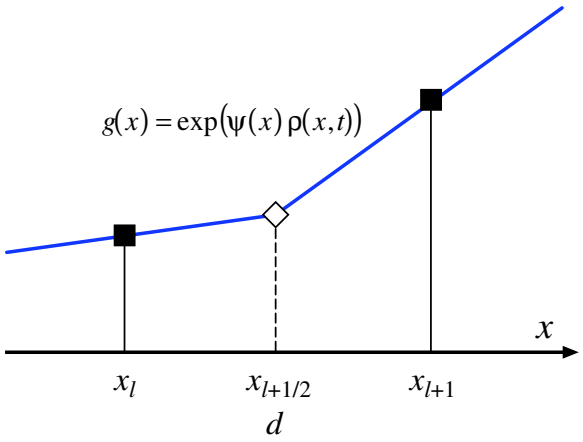

FIGURE 2. Function $g(x) = \exp(\psi(x)\rho(x,t))$ is continuous everywhere but its derivative is discontinuous at $x = x_{l+1/2}$.

Case 1: discontinuity is at a half grid point

We first consider the case where the discontinuity is at a half grid point (that is, $d = x_{l+1/2}$). This is also the most frequently occurring case in modeling molecular motors. Since the discontinuity is at a border separating two cells, in the numerical method, we use $\psi_j^c = \psi_j$ for all cells. Recall condition (29), which implies that function $g(x) = \exp(\psi(x)\rho(x,t))$ is continuous everywhere. But its derivative $g'(x)$ may be discontinuous at $x = x_{l+1/2}$, as illustrated in Figure 2. Our goal is to find a good numerical approximation for $J(x_{l+1/2}, t)$ using only $g(x_l)$ and $g(x_{l+1})$. For that purpose, we multiply (37) by $\exp(\psi(x))$ and integrate from x_l to x_{l+1}.

$$g(x_l) - g(x_{l+1}) = \int_{x_l}^{x_{l+1}} \exp(\psi(x))J(x,t)dx \tag{40}$$

As required by condition (29), probability flux $J(x,t)$ is continuous at $x = x_{l+1/2}$. Its derivative $\frac{\partial}{\partial x}J(x,t)$ is discontinuous at $x = x_{l+1/2}$. The behavior

of $\frac{\partial}{\partial x}J(x,t)$ is also constrained by condition (29). More specifically, taking derivative with respect to t in (29) yields

$$\exp(\psi(d^+))\frac{\partial}{\partial t}\rho(d^+,t) = \exp(\psi(d^-))\frac{\partial}{\partial t}\rho(d^-,t) \tag{41}$$

Applying differential equation (25) leads to

$$\exp(\psi(d^+))\frac{\partial}{\partial x}J(d^+,t) = \exp(\psi(d^-))\frac{\partial}{\partial x}J(d^-,t) \tag{42}$$

In (40), expanding $J(x,t)$ separately to the left of $x = x_{l+1/2}$ and to the right of $x = x_{l+1/2}$, then using relation (42) and dividing by h, we obtain

$$\begin{aligned}
&\frac{g(x_l) - g(x_{l+1})}{h} = \frac{1}{h}\int_{x_l}^{x_{l+1}} \exp(\psi(x))J(x,t)dx \\
&= J(x_{l+1/2},t)\frac{1}{h}\int_{x_l}^{x_{l+1}} \exp(\psi(x))dx \\
&\quad + \frac{1}{h}\int_{x_l}^{x_{l+1}} \exp(\psi(x))[J(x,t) - J(x_{l+1/2},t)]dx \\
&= J(x_{l+1/2},t)\frac{1}{h}\int_{x_l}^{x_{l+1}} \exp(\psi(x))dx \\
&\quad + \left(\exp(\psi(d^+))\frac{\partial}{\partial x}J(d^+,t) - \exp(\psi(d^-))\frac{\partial}{\partial x}J(d^-,t)\right)\frac{1}{h}\int_0^{h/2} xdx + O(h^2) \\
&= J(x_{l+1/2},t)\exp(\psi^b_{l+1/2}) + O(h^2)
\end{aligned} \tag{43}$$

where $\psi^b_{j+1/2}$ is the free energy of border $x_{j+1/2}$ defined as

$$\psi^b_{j+1/2} = \log\left(\frac{1}{h}\int_{x_j}^{x_{j+1}} \exp(\psi(s))ds\right) \tag{44}$$

Of course, the derivation in (43) is also valid for $x_{j+1/2}$ where $\psi(x)$ is continuous. Therefore, the probability flux at $x_{j+1/2}$, no matter whether or not $x_{j+1/2}$ is a discontinuity, is calculated approximately by dividing (43) by $\exp(\psi^b_{l+1/2})$, and the result is used as the numerical flux.

$$\begin{aligned}
J(x_{J+1/2},t) &= \frac{g(x_l) - g(x_{l+1})}{h\ \exp(\psi^b_{l+1/2})} + O(h^2) \\
&\approx \frac{1}{h^2}\cdot\frac{\exp(\psi^c_j)p_j(t) - \exp(\psi^c_{j+1})p_{j+1}(t)}{\exp(\psi^b_{j+1/2})} \equiv J_{j+1/2}(t)
\end{aligned} \tag{45}$$

Here we use $\psi^b_{j+1/2}$ to denote the free energy of border $x_{j+1/2}$ to distinguish it from $\psi_{j+1/2}$, which by convention is defined as $\psi_{j+1/2} = \psi(x_{j+1/2})$. The exact value of the integral in definition (44) is usually difficult to evaluate. For $j \neq l$, $\psi(x)$ is smooth in (x_j, x_{j+1}), and we have $\psi^b_{j+1/2} = \log\left((\exp(\psi_j) + \exp(\psi_{j+1}))/2\right) + O(h^2)$. At $j = l$, $\psi(x)$ is discontinuous

at $x = x_{l+1/2}$ and we only have $\psi^b_{l+1/2} = \log\left((\exp(\psi_l) + \exp(\psi_{l+1}))/2\right) + O(h)$. Although we can have a higher order approximation for $\psi^b_{l+1/2}$ by using $\psi(d^-)$ and $\psi(d^+)$,

(46)
$$\psi^b_{l+1/2} = \log\left(\frac{\exp(\psi_l) + \exp(\psi(d^-)) + \exp(\psi_{l+1}) + \exp(\psi(d^+)}{4}\right) + O(h^2)$$

we will see later that the higher order approximation for $\psi^b_{l+1/2}$ is not essential for achieving the second order accuracy of the numerical method. So in the numerical method, for all cell x_j and for all borders $x_{j+1/2}$, we use

(47)
$$\begin{aligned} \psi^c_j &= \psi_j \\ \psi^b_{j+1/2} &= \log\left(\frac{\exp(\psi_j) + \exp(\psi_{j+1})}{2}\right) \end{aligned}$$

The biggest advantage of (47) is that we do not need to distinguish and identify the location of discontinuity. We use the same formula for all cells and all borders. Writing the numerical flux (45) into the form of a jump process, we have

(48)
$$\begin{aligned} J_{j+1/2}(t) &= F_{j+1/2}p_j(t) - B_{j+1/2}p_{j+1}(t) \\ F_{j+1/2} &= \frac{1}{h^2}\cdot\frac{2\exp(\psi_j)}{\exp(\psi_j) + \exp(\psi_{j+1})} \\ B_{j+1/2} &= \frac{1}{h^2}\cdot\frac{2\exp(\psi_{j+1})}{\exp(\psi_j) + \exp(\psi_{j+1})} \end{aligned}$$

(48) is the improved WPE method for the case where the discontinuity of potential $\psi(x)$ is at a half grid point. We will show in the subsequent sections both analytically and numerically that in this case the improved WPE method has second order accuracy in L^1 norm, L^2 norm and L^∞ norm.

Case 2: the general case

We now consider the general case where the discontinuity at $x = d$ is not restricted to half grid points. Suppose $d \in [x_{l-1/2}, x_{l+1/2})$, as shown in Figure 3. Let $\alpha = \frac{1}{h}(d - x_l)$. α can be positive, which means the discontinuity is to the right of x_l, or α can be negative, which means the discontinuity is to the left of x_l. Our goal is to find first order approximations for free energy of the cell affected by the discontinuity (ψ^c_l) and free energies of the borders affected by the discontinuity ($\psi^b_{l+1/2}$ and $\psi^b_{l-1/2}$). Numerical simulations show that second order approximations for ψ^c_l and $\psi^b_{l\pm1/2}$ are not required for achieving the second order accuracy in L^1 norm. Numerical simulations also show that when the discontinuity is not at a half grid point, second order approximations for ψ^c_l and $\psi^b_{l\pm1/2}$ will not help achieving second order accuracy in L^2 norm and L^∞ norm. As a matter of fact, it is still an open problem that whether or not a linear, 3-point method in conservation form

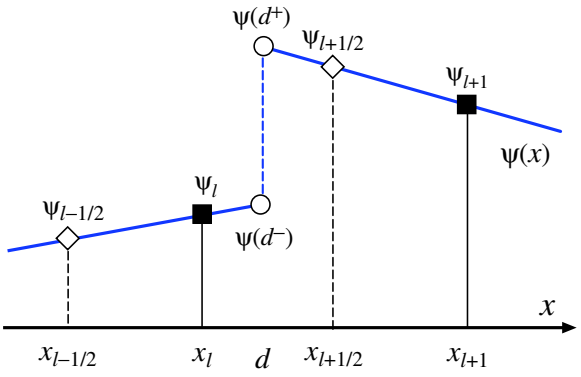

FIGURE 3. Potential $\psi(x)$ is discontinuous at $x = d$ in subinterval $[x_{l-1/2}, x_{l+1/2})$.

can achieve second order accuracy in L^2 norm and L^∞ norm when the discontinuity is not at a half grid point. We also try to make the formulas for case 2 consistent with those of case 1 in the sense that if the discontinuity falls at a half grid point then the formulas of case 2 automatically reduce to those of case 1. For the free energy of cell x_l, we use

$$\begin{aligned}\psi_l^c &= -\log\left(\frac{1}{h}\int_{x_{l-1/2}}^{x_{l+1/2}} \exp(-\psi(s))ds\right) \\ &\approx -\log\left(2|\alpha|\exp(-\psi_l) + (1-2|\alpha|)\frac{\exp(-\psi_{l-1/2}) + \exp(-\psi_{l+1/2})}{2}\right) \qquad (49)\end{aligned}$$

When $|\alpha| = 0.5$, case 2 reduces to case 1 and (49) reduces to (47). For the free energies of borders affected by the discontinuity, we use

$$\begin{aligned}\psi_{l-1/2}^b &= \log\left(\frac{1}{h}\int_{x_{l-1}}^{x_l} \exp(\psi(s))ds\right) \\ &\approx \log\left(|\alpha|(\exp(\psi_{l-1}) + \exp(\psi_l)) + (1-2|\alpha|)\exp(\psi_{l-1/2})\right) \qquad (50) \\ \psi_{l+1/2}^b &= \log\left(\frac{1}{h}\int_{x_l}^{x_{l+1}} \exp(\psi(s))ds\right) \\ &\approx \log\left(|\alpha|(\exp(\psi_l) + \exp(\psi_{l+1})) + (1-2|\alpha|)\exp(\psi_{l+1/2})\right) \qquad (51)\end{aligned}$$

Again, when $|\alpha| = 0.5$, case 2 reduces to case 1, and (50) and (51) reduce to (47). For case 2, writing the numerical flux (45) into the form of a jump

process, we have

$$
\begin{aligned}
J_{j+1/2}(t) &= F_{j+1/2}p_j(t) - B_{j+1/2}p_{j+1}(t) \\
F_{j+1/2} &= \frac{1}{h^2}\cdot\frac{\exp(\psi_j^c)}{\exp(\psi_{j+1/2}^b)} \\
B_{j+1/2} &= \frac{1}{h^2}\cdot\frac{\exp(\psi_{j+1}^c)}{\exp(\psi_{j+1/2}^b)}
\end{aligned}
\tag{52}
$$

where ψ_j^c and $\psi_{j+1/2}^b$ are calculated according to

$$
\begin{aligned}
\psi_j^c &= \begin{cases} \text{Use Eq}(49), & j = l \\ \text{Use Eq}(47), & \text{otherwise} \end{cases} \\
\psi_{j+1/2}^b &= \begin{cases} \text{Use Eq}(50) \text{ or } (51), & j = l-1 \text{ or } j = l \\ \text{Use Eq}(47), & \text{otherwise} \end{cases}
\end{aligned}
$$

(52) is the improved WPE method for the general case where the discontinuity is in $[x_{l-1/2}, x_{l+1/2})$. If the discontinuity is at a half grid point, then the improved WPE method has a very simple form given by (48).

4. Consistency of the improved WPE method

In this section, we study the consistency of the improved WPE method for the case where potential $\psi(x)$ is piecewise smooth and has finite number of discontinuities at half grid points. Without loss of generality, we assume that there is only one discontinuity at $x = d = x_{l+1/2}$ We assume that all derivatives of $\psi(x)$ are continuous away from the discontinuity and have finite left limits and right limits at the discontinuity. In other words, we assume that $\psi(x)$ is two nice functions connected by the discontinuity. Below we will show that away from the discontinuity, the truncation error on the exact solution is $O(h^2)$. At the discontinuity, the truncation error on the exact solution is $O(1)$. However, if we perturb the exact solution by a term of the order $O(h^2)$, then the truncation error on the perturbed solution (instead of the exact solution) is $O(h^2)$ everywhere.

The improved WPE method can be written as

$$
\frac{d}{dt}\left(\frac{p_j(t)}{h}\right) = -\frac{1}{h}(J_{j+1/2}(t) - J_{j-1/2}(t))
\tag{53}
$$

where when the discontinuity is at a half grid point the numerical flux is always

$$
J_{j+1/2}(t) = \frac{2}{h}\left(\frac{\exp(\psi_j)\frac{p_j(t)}{h} - \exp(\psi_{j+1})\frac{p_{j+1}(t)}{h}}{\exp(\psi_j) + \exp(\psi_{j+1})}\right)
\tag{54}
$$

Let $\rho(x,t)$ be the exact solution of Fokker-Planck equation (25) subject to conditions (28) and (29). Let $\rho_j(t) = \rho(x_j, t)$. The truncation error is defined as the residual term when the numerical method is applied to the

exact solution. In the improved WPE method, $\frac{p_j(t)}{h}$ approximates $\rho(x_j,t)$. In (54), replacing $\frac{p_j(t)}{h}$ by $\rho_j(t)$, we have

$$\begin{aligned} J_{j+1/2}\{\rho\} &= \frac{2}{h}\left(\frac{\exp(\psi_j)\rho_j(t)-\exp(\psi_{j+1})\rho_{j+1}(t)}{\exp(\psi_j)+\exp(\psi_{j+1})}\right) \\ &= -\frac{1}{\exp(\psi(x))}\frac{\partial[\exp(\psi(x))\rho(x,t)]}{\partial x}\bigg|_{x_{j+1/2}} \\ &\quad + \begin{cases} a_2(x_{j+1/2},t)h^2+\cdots, & j<l \\ b_2(x_{j+1/2},t)h^2+\cdots, & j>l \\ c_1(t)h+c_2(t)h^2+\cdots, & j=l \end{cases} \end{aligned} \tag{55}$$

where $a_i(x,t)$, $b_i(x,t)$ and $c_i(t)$ are smooth functions of x and t, consisting various derivatives of $\psi(x)$ and $\rho(x,t)$ on both sides of the discontinuity. Here we use $J_{j+1/2}\{\rho\}$ to denote the numerical flux on the exact solution $\rho(x,t)$. This notation is useful when we examine the numerical flux on a perturbed solution $\tilde{\rho}(x,t)$ later. In (55), we have used the fact that $\exp(\psi(x))\rho(x,t)$ is continuous at $x_{l+1/2}$ and is smooth on both sides of $x_{l+1/2}$. The first order term $c_1(t)h$ is caused by that at the discontinuity $x_{l+1/2}$, $(\exp(\psi_l)+\exp(\psi_{l+1}))/2$ is only a first order approximation for $\frac{1}{h}\int_{x_l}^{x_{l+1}}\exp(\psi(s))ds$, as we discussed in the previous section. The truncation error of the improved WPE method is

$$\begin{aligned} E_j\{\rho\} &= \frac{d\rho_j(t)}{dt}+\frac{1}{h}\left(J_{j+1/2}\{\rho\})-J_{j-1/2}\{\rho\})\right) \\ &= \frac{\partial\rho(x,t)}{\partial t}\bigg|_{x_j}+\frac{\partial J(x,t)}{\partial x}\bigg|_{x_j}+O(h^2) \\ &\quad + \begin{cases} O(h^2), & j<l \\ O(h^2), & j>l+1 \\ c_1(t)+\tilde{a}_2(t)h+\cdots, & j=l \\ -c_1(t)+\tilde{b}_2(t)h+\cdots, & j=l+1 \end{cases} \end{aligned} \tag{56}$$

Here we have used the fact that $J(x,t)$ is continuous at $x_{l+1/2}$ and is smooth on both sides of $x_{l+1/2}$, and that since $a_2(x,t)$ is a smooth function, $(a_2(x_{j+1/2},t)-a_2(x_{j-1/2},t))=O(h)$ away from the discontinuity. It is clear that away from the discontinuity the truncation error is $O(h^2)$ while at the discontinuity the truncation error is $O(1)$. The $O(1)$ terms at $j=l$ and $j=l+1$ have the same magnitude and opposite signs. To get rid of the $O(1)$ terms in (56), we need the result of the theorem below.

Theorem 4.1. *Suppose $\{u_j\}$ is periodic in j, satisfies the difference equation*

$$h^2\left[\left(F_{j+1/2}u_j-B_{j+1/2}u_{j+1}\right)-\left(F_{j-1/2}u_{j-1}-B_{j-1/2}u_j\right)\right]=\beta_j \tag{57}$$

and satisfies the constraint $\sum_{j=1}^{M} u_j = 0$ *where* β_j *is*

$$\beta_j = \begin{cases} 1, & j = l \\ -1, & j = l+1 \\ 0, & \textit{otherwise} \end{cases} \tag{58}$$

Then there exists a constant C_ψ*, independent of the numerical grid size, such that*

$$\max_j |u_j| \le C_\psi$$

Proof of Theorem 4.1: Let $\tilde{u}_j = \exp(\psi_j) u_j$. $\tilde{u}_j$ satisfies the difference equation

$$\frac{2\left(\tilde{u}_j - \tilde{u}_{j+1}\right)}{\exp(\psi_j) + \exp(\psi_{j+1})} - \frac{2\left(\tilde{u}_{j-1} - \tilde{u}_j\right)}{\exp(\psi_{j-1}) + \exp(\psi_j)} = \beta_j \tag{59}$$

We construct $\tilde{u}_j$ starting at $j = l+1$ with

$$\tilde{u}_{l+1} = -c_1 \quad \text{and} \quad \frac{1}{h}\,\frac{2\left(\tilde{u}_{l+1} - \tilde{u}_{l+2}\right)}{\exp(\psi_{l+1}) + \exp(\psi_{l+2})} = -c_2$$

where c_1 and c_2 are two coefficients to be determined. Since β_j given in (58) satisfies $\beta_j = 0$ for $j = l+2, \ldots, M+l-1$, we immediately obtain that

$$\frac{1}{h}\,\frac{2\left(\tilde{u}_j - \tilde{u}_{j+1}\right)}{\exp(\psi_j) + \exp(\psi_{j+1})} = -c_2, \quad \text{for } \; j = l+2, \ldots, M+l-1$$

This allows us to write $\tilde{u}_{j+1}$ in terms of $\tilde{u}_j$.

$$\tilde{u}_{j+1} = \tilde{u}_j + c_2 \frac{\exp(\psi_j) + \exp(\psi_{j+1})}{2} h$$

Summing from $j = l+1$ to $j = i-1$, we get

$$\tilde{u}_i = -c_1 + c_2 \sum_{j=l+1}^{i-1} \frac{\exp(\psi_j) + \exp(\psi_{j+1})}{2} h, \quad \text{for } \; i = l+2, \ldots, M+l \tag{60}$$

For $\{\tilde{u}_j\}$ to satisfy (59), it needs to satisfy

$$\frac{2\left(\tilde{u}_l - \tilde{u}_{l+1}\right)}{\exp(\psi_l) + \exp(\psi_{l+1})} = 1 - c_2 h \tag{61}$$

Using that $\tilde{u}_j$ is periodic and substituting (60) into (61) yields an equation for c_2:

$$c_2 \frac{2}{\exp(\psi_l) + \exp(\psi_{l+1})} \sum_{j=l+1}^{M+l-1} \frac{\exp(\psi_j) + \exp(\psi_{j+1})}{2} h = 1 - c_2 h \tag{62}$$

Solving for c_2, we obtain

$$c_2 = \left[h + \frac{2}{\exp(\psi_l) + \exp(\psi_{l+1})} \sum_{j=l+1}^{M+l-1} \frac{\exp(\psi_j) + \exp(\psi_{j+1})}{2} h \right]^{-1} \tag{63}$$

The sum in (63) is approximately an integral

$$\sum_{j=l+1}^{M+l-1} \frac{\exp(\psi_j) + \exp(\psi_{j+1})}{2} h = \int_0^L \exp(\psi(x)) dx + O(h)$$

Substituting this result into (63), we have

$$c_2 = \left[\frac{2}{\exp(\psi_l) + \exp(\psi_{l+1})} \int_0^L \exp(\psi(x)) dx \right]^{-1} + O(h) \tag{64}$$

Thus, for h small enough, c_2 is positive and bounded. c_1 is determined by the constraint $\sum_{j=1}^{M} \exp(-\psi_j)\tilde{u}_j = 0$. Notice that $\tilde{u}_j$, given in (60), is monotonically increasing for $j = l+1, \ldots, M+l$. If $c_1 = 0$, then we have $\tilde{u}_{l+1} = 0$ and $\tilde{u}_j > 0$ for $j = l+2, \ldots, M+l$, which leads to $\sum_{j=1}^{M} e^{-\psi_j} \tilde{u}_j > 0$. Now we select c_1 to make $\tilde{u}_{M+l} = 0$:

$$\hat{c}_1 = c_2 \sum_{j=l+1}^{M+l-1} \frac{\exp(\psi_j) + \exp(\psi_{j+1})}{2} h = \frac{\exp(\psi_l) + \exp(\psi_{l+1})}{2} + O(h)$$

In this case, we have $\tilde{u}_{M+l} = 0$ and $\tilde{u}_j < 0$ for $j = l+1, \ldots, M+l-1$, which leads to $\sum_{j=1}^{M} e^{-\psi_j} \tilde{u}_j < 0$. The value of c_1 that satisfies the constraint $\sum_{j=1}^{M} \exp(-\psi_j)\tilde{u}_j = 0$ must be between 0 and $\hat{c}_1$. Thus, for h small enough, c_1 is positive and bounded.

$$0 < c_1 < \frac{\exp(\psi_l) + \exp(\psi_{l+1})}{2} + O(h) \tag{65}$$

Substituting (64) and (65) into (60), we conclude that

$$\max_j |\tilde{u}_j| \leq \frac{\exp(\psi_l) + \exp(\psi_{l+1})}{2} + O(h) \tag{66}$$

which leads directly to the conclusion of Theorem 4.1.

Periodic boundary condition is the most frequently used boundary condition in modeling molecular motors. The result of Theorem 4.1 is also valid for other boundary conditions. To get rid of the $O(h)$ terms in (56), we need the result of another theorem.

Theorem 4.2. *Suppose $\{v_j\}$ is periodic in j, satisfies the difference equation*

$$h^2 \left[\left(F_{j+1/2} v_j - B_{j+1/2} v_{j+1} \right) - \left(F_{j-1/2} v_{j-1} - B_{j-1/2} v_j \right) \right] = \eta_j \tag{67}$$

and satisfies the constraint $\sum_{j=1}^{M} v_j = 0$ *where* η_j *is*

$$\eta_j = \begin{cases} h, & j = l \\ h, & j = l+1 \\ -2h/(M-2), & \textit{otherwise} \end{cases} \tag{68}$$

Then there exists a constant C_ψ*, independent of the numerical grid size, such that*

$$\max_j |v_j| \le C_\psi$$

Proof of Theorem 4.2 is similar to that of Theorem 4.1 and is skipped.

Now let us consider a perturbed solution on grid points:

$$\tilde{\rho}_j(t) = \rho_j(t) + \left(c_1(t) + \frac{\tilde{a}_2(t) - \tilde{b}_2(t)}{2} h \right) h^2 u_j + \frac{\tilde{a}_2(t) + \tilde{b}_2(t)}{2} h^2 v_j \tag{69}$$

where $\{u_j\}$ and $\{v_j\}$ are given by Theorem 4.1 and Theorem 4.2 above. In a bounded region, the distance between the exact solution and the perturbed solution given by (69) is $O(h^2)$ in any norm. Substituting the perturbed solution into (55) and (56), we obtain

$$E_j\{\tilde{\rho}\} = \frac{d\tilde{\rho}_j(t)}{dt} + \frac{1}{h}\left(J_{j+1/2}\{\tilde{\rho}\}) - J_{j-1/2}\{\tilde{\rho}\}) \right) + O(h^2) \tag{70}$$

which means that for the improved WPE method, the truncation error on the perturbed solution is $O(h^2)$, and the perturbed solution is only $O(h^2)$ away from the exact solution.

After the spatial discretization, Fokker-Planck equation (25) is approximated by the master equation of the jump process (36) (an ODE system). For the time integration of the improved WPE method, we propose to use a two stage second order diagonally implicit Runge Kutta (DIRK) method [15]. Suppose the ODE system we are solving has the form $\vec{y'} = \vec{f}(\vec{y}, t)$. The second order DIRK method we use is

$$\begin{aligned} \vec{k}_1 &= \Delta t \cdot \vec{f}(\vec{y}_n + a\ \vec{k}_1, t_n + a\Delta t) \\ \vec{k}_2 &= \Delta t \cdot \vec{f}(\vec{y}_n + (1-a)\vec{k}_1 + a\ \vec{k}_2, t_n + \Delta t) \\ \vec{y}_{n+1} &= \vec{y}_n + (1-a)\vec{k}_1 + a\ \vec{k}_2 \end{aligned} \tag{71}$$

where $a = 1 - 1/\sqrt{2}$. In comparison with the frequently used Crank-Nicolson method (trapezoidal method) [24], the DIRK method is A-stable and L-stable while the Crank-Nicolson method is only A-stable. Although an A-stable method usually can prevent high wave number oscillations from growing it is not very effective at suppressing high wave number oscillations once they appear. On the other hand, an L-stable method is capable of suppressing high wave number oscillations. Here it is important to notice

that using an L-stable DIRK method for time integration is consistent with the underlying diffusive property of the physical process.

The stability of the WPE method has been proved in [12] and the proof can be extended to the improved WPE method. Once we have both the stability and the consistency, the convergence of the improved WPE method follows from the Lax equivalence theorem [22].

5. Numerical results and discussions

In this section, we compare the performance of the WPE method (21) and the improved WPE method (52) on a model problem with discontinuous potential. In the model problem, we take $k_BT = 1$, $D = 1$ and $L = 1$, and select the tilted periodic potential

$$\psi(x) = \begin{cases} 6 - 6\sin\left(\frac{\pi}{2}(x+1-d)\right), & 0 < x < d \\ 3 - 6\sin\left(\frac{\pi}{2}(x-d)\right), & d < x < 1 \end{cases} \tag{72}$$

Three periods of this tilted periodic potential is shown in Figure 4. In

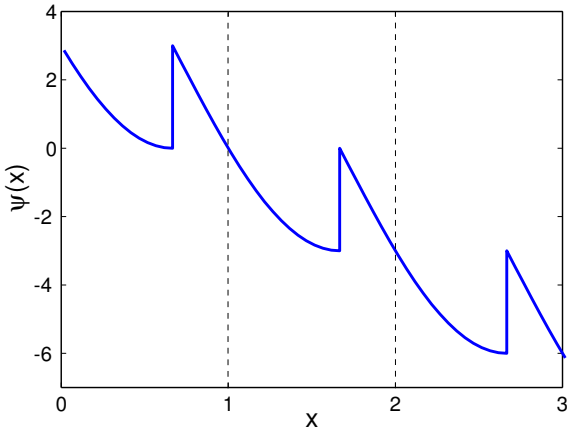

FIGURE 4. Graph of the tilted periodic potential given in (72).

each period, the potential has a discontinuity of magnitude 3 at $x = d$. We compare the performance of the WPE method (21) and the improved WPE method (52) in predicting the steady state probability density, predicting the average velocity and predicting the effective diffusion of the motor [21]. We run simulations with spatial step $h = 2^{-m}$ where m is a positive integer. In simulations below we will test two cases: $d = 3/4$ and $d = 2/3$. When $d = 3/4$, the discontinuity is always at a half grid point. When $d = 2/3$, the discontinuity is never at a half grid point.

We define the error as the difference between the numerical solution obtained with a finite value of h and the exact solution. We estimate the error

as follows. Suppose $\vec{p}(h)$ is the numerical solution obtained with spatial step h. The error in $\vec{p}(h)$ is estimated as

$$\text{error}(h) \approx C_p \left\| \vec{p}(h) - \vec{p}\left(\frac{h}{2}\right) \right\| \tag{73}$$

where C_p is a constant depending on the order of the method. For methods of first order or higher, C_p is between 1 and 2. Here we simply use $C_p = 1$. Thus, for methods of first order or higher, in the worst case scenario, we underestimate the error by a factor of 2.

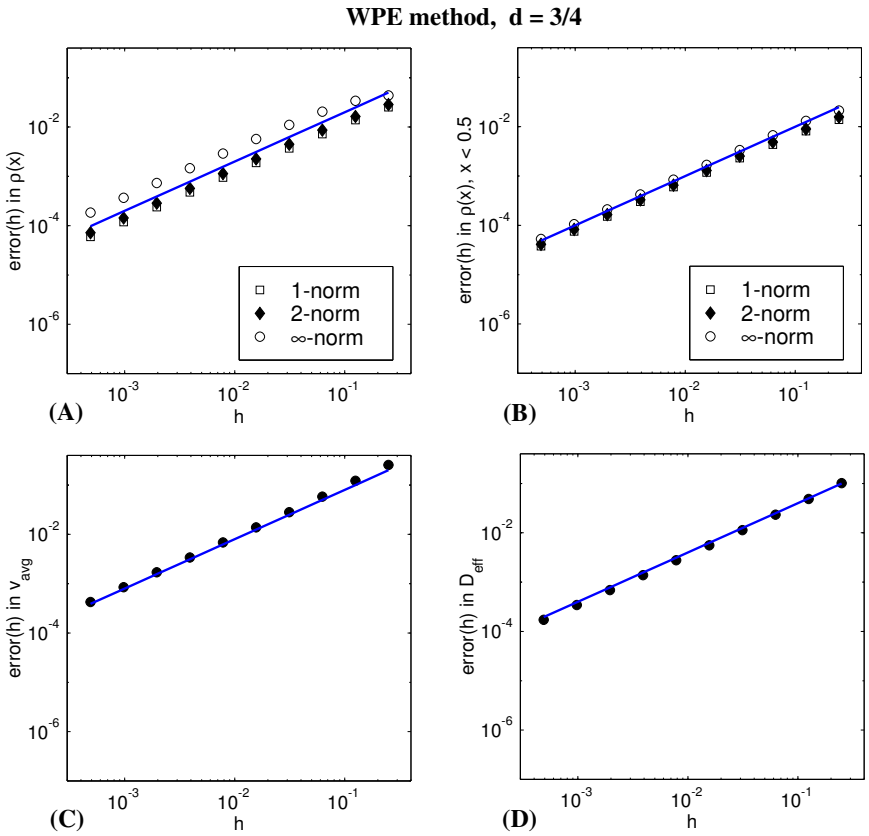

FIGURE 5. Estimated errors of the WPE method for the case of $d = 3/4$. (A) error in the probability density $\rho(x)$, (B) error in the probability density $\rho(x)$ away from the discontinuity, (C) error in the average velocity, and (D) error in the effective diffusion. In four panels, solid lines are plots of $y = c_1 h$ to show the order of convergence. It is clear that when the potential has a discontinuity at a half grid point, the WPE method has first order accuracy.

Figure 5 shows the estimated errors of the WPE method for the case of $d = 3/4$. As expected, everything predicted by the WPE method is first

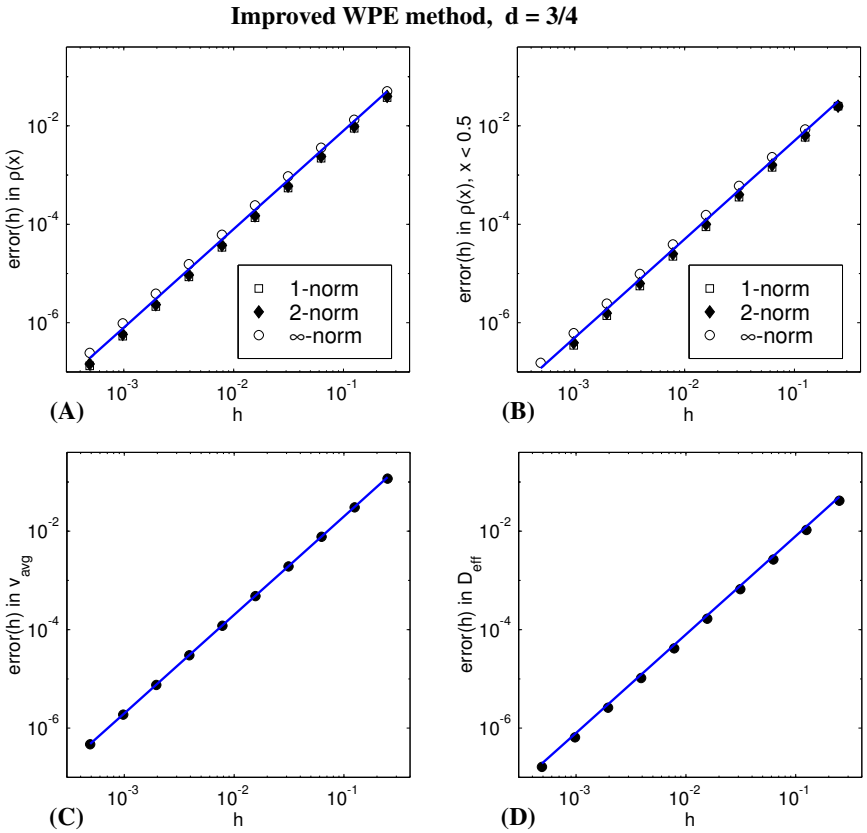

FIGURE 6. Estimated errors of the improved WPE method for the case of $d = 3/4$. (A) error in the probability density $\rho(x)$, (B) error in the probability density $\rho(x)$ away from the discontinuity, (C) error in the average velocity, and (D) error in the effective diffusion. In four panels, solid lines are plots of $y = c_2 h^2$ to show the order of convergence. It is clear that when the potential has a discontinuity at a half grid point, the improved WPE method has second order accuracy.

order when the potential has a discontinuity at a half grid point. Figure 6 shows the estimated errors of the improved WPE method for the case of $d = 3/4$. When the discontinuity of the potential is at a half grid point, everything predicted by the improved WPE method is second order.

The success of the error estimation method (73) depends on that the coefficient of the leading term in error is a constant when the step size is reduced. This is usually not a problem when the discontinuity is at a half grid point. However, when the discontinuity is not at a half grid point, this is generally not true and the error estimation method (73) is no longer valid. So for the case of $d = 2/3$, we need to find another way to estimate the

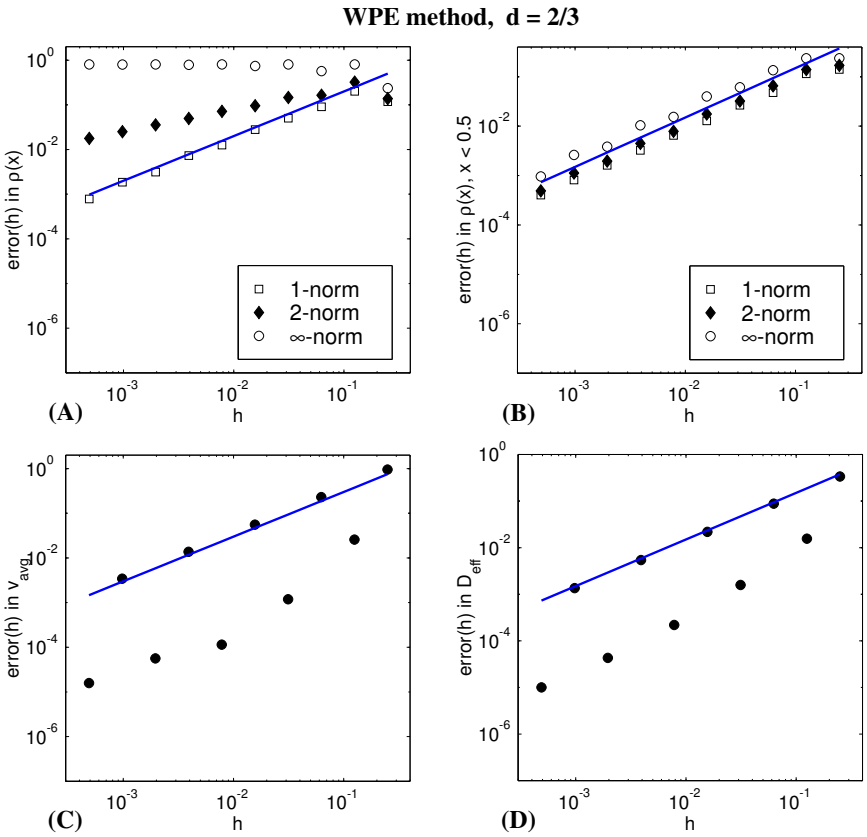

FIGURE 7. Estimated errors of the WPE method for the case of $d = 2/3$. (A) error in the probability density $\rho(x)$, (B) error in the probability density $\rho(x)$ away from the discontinuity, (C) error in the average velocity, and (D) error in the effective diffusion. In four panels, solid lines are plots of $y = c_1 h$ to show the order of convergence. It is clear that when the potential has a discontinuity not at a half grid point, the WPE method has first order accuracy in L^1 norm.

error. For the case of $d = 2/3$, we use a very accurate numerical solution as the exact solution in calculating the error. More specifically, we use $h = 1/6144$ (instead of $h = 2^{-m}$) so $d = 2/3$ is at a half grid point, and we use the improved WPE method to find a very accurate numerical solution. The improved WPE method has been shown in Figure 6 to yield second order accuracy when the discontinuity is at a half grid point.

Figure 7 shows the estimated errors of the WPE method for the case of $d = 2/3$. The WPE method has first order accuracy in L^1 norm when the potential has a discontinuity not at a half grid point. At the discontinuity, the WPE method does not converge in L^∞ norm. Figure 8 shows the estimated errors of the improved WPE method for the case of $d = 2/3$. Even

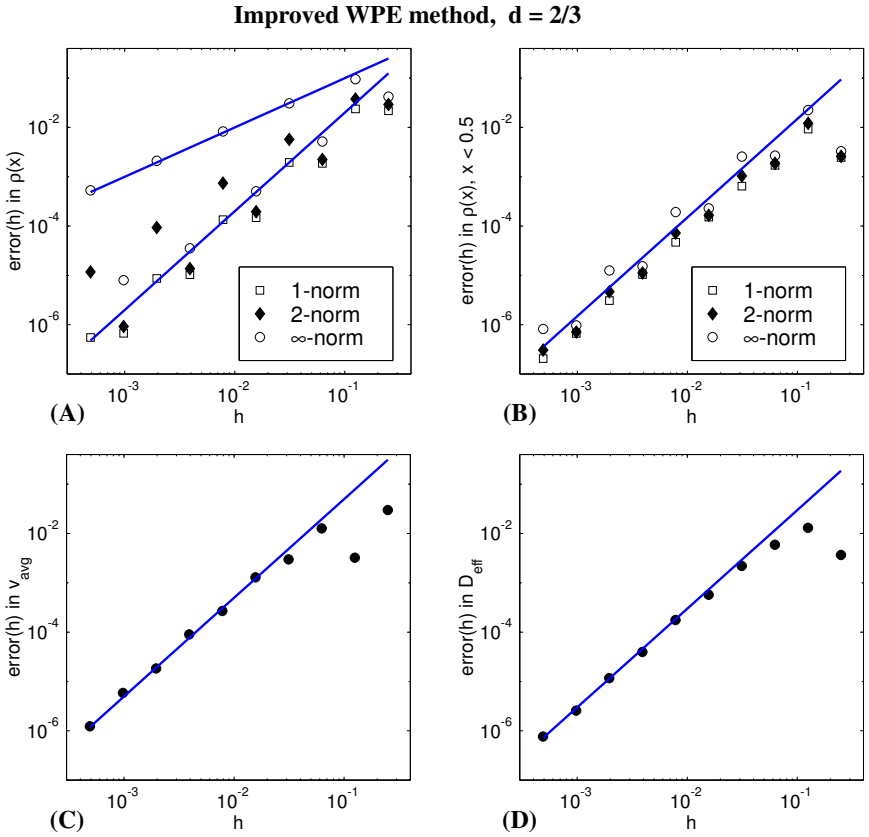

FIGURE 8. Estimated errors of the improved WPE method for the case of $d = 2/3$. (A) error in the probability density $\rho(x)$, (B) error in the probability density $\rho(x)$ away from the discontinuity, (C) error in the average velocity, and (D) error in the effective diffusion. In four panels, solid lines are plots of $y = c_1 h^2$ and $y = c_2 h^2$ to show the order of convergence. It is clear that even when the potential has a discontinuity not at a half grid point, the improved WPE method has second order accuracy in L^1 norm, first order accuracy in L^∞ norm and second order accuracy for average velocity and effective diffusion.

when the potential has a discontinuity not at a half grid point, the improved WPE method has second order accuracy for the average velocity, for the effective diffusion, and for the probability density in all norms away from the discontinuity. Near the discontinuity, the improved WPE method has second order accuracy in L^1 norm and first order accuracy in L^∞ norm. From the robust performance shown in Figures 6 and 8, it is clear that the improved WPE method (52) can handle discontinuous potentials and yield

second order accuracy for all interesting quantities in modeling molecular motors even if the discontinuity is not at a half grid point.

In conclusions, we have designed an improved version of the WPE method for solving discontinuous Fokker-Planck equations. The improved WPE method has two advantages: 1) it preserves the second order accuracy when the potential is discontinuous and even if the discontinuity is not at a half grid point, and 2) it eliminates the numerical singularity in the jump rates of the WPE method. Also the improved WPE method is as simple and easy to implement as the WPE method.

Acknowledgments

The author would like to thank Charles S. Peskin and Timothy C. Elston for helpful discussions. This work was partially supported by NSF.

References

[1] H. C. Berg, *Random walks in biology*, Princeton University Press, Princeton, N.J., 1993.

[2] J. Abrahams, A. Leslie, R. Lutter, and J. Walker, *Structure at 2.8A resolution of F1-ATPase from bovine heart mitochondria*, Nature, **370**, 621-628, 1994.

[3] H. Noji, R. Yasuda, M. Yoshida, and K. Kinosita, *Direct observation of the rotation of F1-ATPase*, Nature, **386**, 299-302, 1997.

[4] H. Wang and G. Oster, *Energy transduction in the F1 motor of ATP synthase*, Nature, **396**, 279-282, 1998.

[5] C. Coppin, D. Pierce, L. Hsu, and R. Vale, *The Load Dependence of Kinesin's Mechanical Cycle*, Proc. Natl. Acad. Sci. USA, **94**, 8539-8544, 1997.

[6] K. Visscher, M. Schnitzer, and S. Block, *Single kinesin molecules studied with a molecular force clamp* Nature, **400**, 184-189, 1999.

[7] J. Prost, J. Chauwin, L. Peliti, and A. Ajdari, *Asymmetric Pumping of Particles*, Phys. Rev. Lett., **72**, 2652-2655, 1994.

[8] R. Astumian, *Thermodynamics and Kinetics of a Brownian Motor*, Science, **276**, 917-922, 1997.

[9] T. Elston, H. Wang, and G. Oster, *Energy transduction in ATP synthase*, Nature, **391**, 510-514, 1998.

[10] H. Wang, *Mathematical theory of molecular motors and a new approach for uncovering motor mechanism*, IEE Proceedings Nanobiotechnology, **150**, 127-133, 2003.

[11] H. Wang, *Chemical and mechanical efficiencies of molecular motors and implications for motor mechanisms*, Journal of Physics: Condensed Matter, **17**, S3997-S4014, 2005.

[12] H. Wang, *Convergence of a numerical method for solving discontinuous Fokker-Planck equations*, Preprint, 2005.

[13] F. Reif, *Fundamentals of Statistical and Thermal Physics*, New York, McGraw-Hill, 1985.

[14] H. Risken, *The FokkerPlanck Equation*, 2nd edn., Berlin, Springer, 1989

[15] E. Hairer, *Solving Ordinary Differential Equations II : Stiff and Differential-Algebraic Problems*, 2nd edn., Vol 2, Springer Verlag, 1996

[16] A. Einstein, *Investigation on the Theory of the Brownian Motion*. New York, Dover, 1956.

[17] S. R. de Groot and P. Mazur, *Non-Equilibrium Thermodynamics*. New York, Dover, 1984.
[18] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II*. Berlin, Springer, 1995.
[19] C. Peskin, G. Odell, and G. Oster, *Cellular motions and thermal fluctuations: The Brownian ratchet*, Biophys. J., **65**, 316-324, 1993.
[20] T. Elston and C. Peskin, *The role of protein flexibility in molecular motor function: Coupled diffusion in a tilted periodic potential*, SIAM J. Appl. Math., **60**, 842867, 2000.
[21] H. Wang, C. Peskin and T. Elston, *A Robust numerical algorithm for studying biomolecular transport processes*, J. Theo. Biol., **221**, 491-511, 2003.
[22] R. D. Richtmyer, and K. W. Morton, *Difference methods for initial value problems*, Wiley-Interscience, New York, 1967.
[23] T. Elston, and C. Doering, *Numerical and analytical studies of nonequilibrium fluctuation induced transport processes*, J. Stat. Phys., **83**, 359-383, 1996.
[24] M. L. Juncosa, and D. Young, *On the Crank-Nicolson procedure for solving parabolic partial differential equations*, Proc. Cambridge Philos. Soc., **53**, 448-461, 1957.
[25] R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison-Wesley, Reading, MA, 1963.

Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA 95064, USA

E-mail: hongwang@ams.ucsc.edu

URL: http://www.soe.ucsc.edu/∼hongwang/