Available online at www.sciencedirect.com

ScienceDirect

Journal of Theoretical Biology 242 (2006) 908–921

Journal of Theoretical Biology

www.elsevier.com/locate/yjtbi

# Motor potential profile and a robust method for extracting it from time series of motor positions

Hongyun Wang

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

Received 27 December 2005; received in revised form 4 April 2006; accepted 5 April 2006
Available online 28 April 2006

## Abstract

Molecular motors are small, and, as a result, motor operation is dominated by high-viscous friction and large thermal fluctuations from the surrounding fluid environment. The small size has hindered, in many ways, the studies of physical mechanisms of molecular motors. For a macroscopic motor, it is possible to observe/record experimentally the internal operation details of the motor. This is not yet possible for molecular motors. The chemical reaction in a molecular motor has many occupancy states, each having a different effect on the motor motion. The overall effect of the chemical reaction on the motor motion can be characterized by the motor potential profile. The potential profile reveals how the motor force changes with position in a motor step, which may lead to insights into how the chemical reaction is coupled to force generation. In this article, we propose a mathematical formulation and a robust method for constructing motor potential profiles from time series of motor positions measured in single molecule experiments. Numerical examples based on simulated data are shown to demonstrate the method. Interestingly, it is the small size of molecular motors (negligible inertia) that makes it possible to recover the potential profile from time series of motor positions. For a macroscopic motor, the variation of driving force within a cycle is smoothed out by the large inertia.

*Keywords:* Molecular motors; Stochastic processes; Fokker–Planck equations; Time series analysis

## 1. Introduction

Molecular motors are very small and operate in a viscous fluid environment (water). Consequently, the motor motion is dominated by high viscous friction and large thermal fluctuations (Berg, 1993). The small size of molecular motors has hindered, in many ways, the studies of motor mechanisms. It is not yet possible to observe/record the internal motion of a single molecule motor. Nor is it possible to observe/record the chemical occupancy state of a single molecule motor. Because of the small size, molecular motors have several features that distinguish themselves from macroscopic motors. The most prominent feature of molecular motors is that the time scale of inertia is much smaller than that of chemical reaction cycle in the motor. On the time scale of motor operation, the effect of inertia is negligible. Another feature of molecular motors is that the instantaneous velocity, caused by the bombardments of surrounding fluid molecules, changes drastically over the time scale of motor operation, and is typically much larger than the average velocity of the motor. In comparison, for macroscopic motors, the time scale of inertia is much larger than that of reaction cycle, and the instantaneous velocity is almost constant over the time scale of reaction cycle and is almost the same as the average velocity. These peculiar features suggest that we may be able to exploit some advantages of molecular motors that are not present in macroscopic motors.

The current experimental technologies allow us to measure forces and motions of a single motor to the precision of piconewtons and nanometers (Abbondanzieri et al., 2005; Hunt et al., 1994; Schnitzer and Block, 1997; Itoh et al., 2004; Hirono-Hara et al., 2005). Time series of motor positions have been measured for various molecular motors at various mechanical loads and chemical concentrations (Yasuda et al., 1998; Visscher et al., 1999;

*E-mail address:* hongwang@ams.ucsc.edu.

0022-5193/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2006.04.005

Sambongi et al., 1999; Block et al., 2003). In the past, only the average velocity and sometimes the randomness parameter of the motor were extracted from the measured time series of motor positions (Visscher et al., 1999; Samuel and Berg, 1995). In Wang (2003), we proposed the concept of motor potential profile. The chemical reaction in a molecular motor has many occupancy states, each having a different effect on the motor motion. The overall effect of the chemical reaction on the motor motion can be characterized by the motor potential profile. At each motor position, the motor force profile is the average motor force over all chemical states at that position weighted by the steady-state probability densities. The motor force profile is a periodic function of motor position. The integral of the motor force profile is the motor potential profile, which is a tilted periodic function. The motor potential profile does not contain all information about motor operation. In particular, it does not capture the full picture of chemical reaction in the motor. The most important reason we study the motor potential profile is that, in principle, it can be recovered from measured time series of motor positions (Wang, 2003). In this sense, the potential profile is a measurable entity (in addition to average velocity and randomness parameter). The potential profile reveals how the motor force changes with position in a motor step, which provides insights into how the chemical reaction is coupled to force generation.

In both macroscopic motors and molecular motors, a unidirectional motion can be produced by generating an active force at the chemical reaction site and using the active force to drive the motor forward. This mechanism of producing a unidirectional motion is called power stroke motor (Wang and Oster, 1998; Oster and Wang, 2000). In molecular motors, a unidirectional motion can also be produced by rectifying thermal fluctuations: if thermal fluctuations in the backward direction are blocked by a free energy barrier, then the motor will be effectively carried forward by thermal fluctuations. This mechanism of producing a unidirectional motion is called Brownian ratchet (Peskin et al., 1993; Elston et al., 1998; Mogilner and Oster, 1999; Astumian, 1997; Reimann, 2002) or information ratchet (Astumian and Derenyi, 1998). In a power stroke motor, the chemical reaction generates an active driving force, which corresponds to a gradually decreasing potential. In a Brownian ratchet, the chemical reaction establishes a free energy barrier, which corresponds to a vertical drop followed by a flat step in free energy. Thus, the potential profile of a Brownian ratchet is a sequence of vertical free energy drops rectifying forward fluctuations. The potential profile of a power stroke motor is a gradually decreasing function of the motor position, generating an active force to drive the motor. Of course, Brownian ratchet and power stroke motor are just two extreme situations. The potential profile of a motor may have both vertical free energy drops and down hill slopes. The motor potential profile is the link between the chemical reaction and the motor motion. We can conceptually divide the motor into two parts: first the chemical reaction generates the potential profile; then the potential profile produces the unidirectional motion.

In this paper, we study the mathematical formulation and method for constructing the motor potential profile from time series of motor positions. Below, we will first introduce the mathematical framework for modeling the continuous motion of molecular motors. Then we will review the motor potential profile proposed in Wang (2003), which is defined by averaging motor force at each position over all chemical states. In principle, the motor potential profile can be recovered in two steps: (1) differentiating/differencing the time series to get the total force of the motor and (2) summing at each motor position to get rid of the fluctuating Brownian force. Most of techniques in time series analysis follow roughly this procedure of differencing a non-stationary time series (position) to obtain a stationary time series (velocity) and then fitting models based on the theory of stationary time series. This allows one to account for the dependency structure in the data (Brockwell and Davis, 1998; Fan and Yao, 2005). In this study, our goal is to recover the motor potential profile and we know (or we assume) the mathematical framework governing the generation of data. The availability of the mathematical framework for the underlying physical process allows us to adopt a more robust mathematical formulation by rewriting the motor potential profile in terms of the steady-state probability flux and the steady-state probability density. The steady-state probability flux is related to the average velocity, which can be calculated reliably from data. After that we focus on how to estimate the probability density and how to estimate the statistical errors associated with the estimated probability density function. To achieve a good balance between spatial resolutions and statistical errors, we use an adaptive spatial resolution based on the data to obtain the optimal result. The statistical errors are estimated, and based on the estimated errors the spatial resolution can be tuned to achieve an optimal result. This is especially important when we have only a limited amount of data. Finally, numerical examples based on simulated data are shown to demonstrate the method.

## 2. Mathematical framework for modeling the continuous motion of molecular motors

In this section, we discuss the mathematical framework for modeling the continuous motion of molecular motors. In many chemical reactions, the mechanical motion involves only a small conformational change against no significant force, and it occurs at a time scale much smaller than that of the reaction cycle. These chemical reactions can be well described by simple kinetic models. In molecular motors, however, the mechanical motion generally involves a large conformational change against a significant conservative load and/or viscous drag, and it occurs over a time scale comparable to that of the reaction cycle. These special properties of molecular motors require that the mechanical motion be modeled explicitly as a continuous motion and be coupled to the chemical reaction. One advantage of modeling the continuous motion of molecular motors is that physical quantities of continuous motion can be incorporated into and discussed in the model. For example, the drag force, the Stokes efficiency (Wang and Oster, 2002) and the mechanical coupling of multiple motors can be discussed in the framework of continuous motion.

In general, the continuous motion of a molecular motor has many degrees of freedom, of which there is a prominent one associated with the unidirectional motion of the motor. For example, the $\gamma$ shaft of the $F_0F_1$ ATP synthase rotates with respect to the $\alpha_3\beta_3$ hexamer (Abrahams et al., 1994; Sabbert et al., 1996; Noji et al., 1997; Wang and Oster, 1998; Menz et al., 2001), a flagellar motor rotates the flagellar filament with respect to the cell body (Block and Berg, 1984; Berg, 2003) and a kinesin dimer moves along a microtubule (Vale, 1986; Howard et al., 1989; Visscher et al., 1999; Coppin et al., 1997). The first stage of modeling the continuous motion is to follow the motor only along the dimension of its unidirectional motion and include the effects of other degrees of freedom in the mean field potential affecting the unidirectional motion (Prost et al., 1994; Julicher et al., 1997; Astumian, 1997; Elston et al., 1998). Let us start by looking at the one-dimensional motion of a small particle in a fluid environment, subject to a potential, $V(x)$, where $x$ is the coordinate along the dimension of unidirectional motion. The particle is driven by the conservative force derived from the potential, by the viscous drag, and by the Brownian force. The stochastic motion of the particle is governed by the Langevin equation with inertia (Newton's second law):

$$m\frac{\mathrm{d}v}{\mathrm{d}t} = \underbrace{-\zeta v}_{\substack{\text{Viscous}\\ \text{drag}}} \ \underbrace{-V'(x)}_{\substack{\text{Force from}\\ \text{potential}}} + \underbrace{\sqrt{2k_BT\zeta}\,\frac{\mathrm{d}W(t)}{\mathrm{d}t}}_{\substack{\text{Brownian}\\ \text{force}}}, \tag{1}$$

where $m$ is the mass and $v$ the velocity of the particle. $W(t)$ is the Weiner process. The viscous drag on the particle, $-\zeta v$, is always opposing the motion where $\zeta$ is called the drag coefficient. The magnitude of the Brownian force is related to the drag coefficient by $\sqrt{2k_BT\zeta}$, which is a result of the fluctuation–dissipation theorem (Reif, 1965). Here $k_B$ is the Boltzmann constant, and $T$ the absolute temperature (Landau et al., 1980).

Because of the small size of molecular motors, motor operation is dominated by high viscous friction and thermal fluctuations from the surrounding fluid environment. In Eq. (1), the evolution of particle has two very different time scales: the very short time scale of the particle forgetting about its current instantaneous velocity (the time scale of inertia) and the relatively long time scale of the particle moving along the potential. In the absence of the potential, we have

$$\frac{\mathrm{d}\langle v(t)\rangle}{\mathrm{d}t} = -\frac{\zeta}{m}\langle v(t)\rangle.$$

The time scale of inertia is $t_0 = m/\zeta$. Suppose the particle is a bead of radius $r$. The drag coefficient, the mass, and the time scale of inertia of the particle are, respectively, given by Berg (1993),

$$\zeta = 6\pi\eta r, \quad m = \frac{4}{3}\pi\rho r^3, \quad t_0 = \frac{m}{\zeta} = \frac{2\rho}{9\eta}r^2,$$

where $\rho$ is the density of the particle and $\eta$ the viscosity of the surrounding fluid. Since the time scale $t_0$ is proportional to the square of the radius, for a small particle, $t_0$ is very small. For a particle of radius 0.5 μm in water, we have $t_0 \approx 56 \times 10^{-9}$ s, which is much smaller than the time scale of motor reaction cycles. If we are only concerned with evolution over time scales much larger than $t_0$, we can safely ignore the effect of inertia and simplify Eq. (1) significantly. We rewrite Eq. (1) as

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -\frac{1}{t_0}\left(v - \left[-\frac{1}{\zeta}V'(x) + \sqrt{2D}\frac{\mathrm{d}W(t)}{\mathrm{d}t}\right]\right), \tag{2}$$

where $D = k_BT/\zeta$ is the diffusion coefficient (Berg, 1993). When $t_0$ is very small, we have approximately

$$v - \left[-\frac{1}{\zeta}V'(x) + \sqrt{2D}\frac{\mathrm{d}W(t)}{\mathrm{d}t}\right] = 0. \tag{3}$$

Writing (3) as a differential equation for $x$ yields

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\frac{1}{\zeta}[-f + \phi'_S(x)] + \sqrt{2D}\frac{\mathrm{d}W(t)}{\mathrm{d}t}. \tag{4}$$

Here for the convenience of discussing motor systems, we have written the potential $V(x)$ explicitly as a constant force $f$ and a periodic potential $\phi_S(x)$, which varies with the current chemical occupancy state (S) of the motor. Eq. (4) is the Langevin equation without inertia governing the stochastic motion of a small particle driven by a constant external force $f$ and a periodic potential $\phi_S(x)$ that is coupled to the chemical reaction.

In molecular motors, the motion is driven by switching among a set of potentials, each corresponding to a chemical occupancy state. In Eq. (4), the periodic potential $\phi_S(x)$ changes as the chemical reaction proceeds in the motor system (Prost et al., 1994; Elston et al., 1998). Suppose the chemical reaction cycle has $N$ occupancy states. For ATPase motors, each catalytic site has four occupancy states (Abrahams et al., 1994; Boyer, 1993, 1997; Weber and Senior, 1997):

$$\begin{aligned}
&[\textbf{Site}] + \text{ATP} \underset{\text{Binding}}{\longleftrightarrow} [\textbf{Site}\cdot\textbf{ATP}] \underset{\text{Hydrolysis}}{\longleftrightarrow} [\textbf{Site}\cdot\textbf{ADP}\cdot\textbf{P}_\textbf{i}] \\
&\qquad \underset{\text{P}_\text{i}\ \text{release}}{\longleftrightarrow} [\textbf{Site}\cdot\textbf{ADP}] + \text{P}_\text{i} \\
&\qquad \underset{\text{ADP release}}{\longleftrightarrow} [\textbf{Site}] + \text{ADP} + \text{P}_\text{i}.
\end{aligned}$$

The interaction between a kinesin dimer and a microtubule has at least two occupancy states: (1) one head bound to the microtubule and (2) both heads bound to the

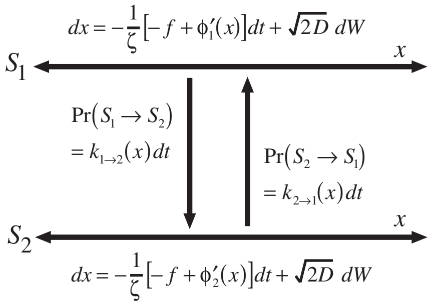
$dx = -\frac{1}{\zeta}[-f + \phi_1'(x)]dt + \sqrt{2D}\, dW$
$S_1$
$x$
$\Pr(S_1 \to S_2) = k_{1\to2}(x)dt$
$\Pr(S_2 \to S_1) = k_{2\to1}(x)dt$
$S_2$
$x$
$dx = -\frac{1}{\zeta}[-f + \phi_2'(x)]dt + \sqrt{2D}\, dW$

Fig. 1. The mechanical motion (horizontal direction) and the chemical transition (vertical direction) in a two-state motor system.

microtubule. Thus, a kinesin dimer moving along a microtubule has, at least, $N = 4 \times 4 \times 2 = 32$ occupancy states (counting the two ATP catalytic sites and the interaction with microtubule). We use a discrete Markov process to model the stochastic evolution of the chemical occupancy state. Let $S(t)$ be the chemical occupancy state of the motor system at time $t$. $S(t)$ takes values in the set $\{1, 2, \ldots, N\}$. $S(t)$ evolves stochastically according to the equation below:

$$\begin{aligned}&\Pr[S(t+\Delta t) = i|S(t) = j] = k_{j\to i}(x)\Delta t + o(\Delta t), \quad i \neq j,\\ &\Pr[S(t+\Delta t) = j|S(t) = j] = 1 - \sum_{i\neq j} k_{j\to i}(x)\Delta t + o(\Delta t), \quad (5)\end{aligned}$$

where $k_{j\to i}(x)$ is the transition rate from occupancy state $j$ to state $i$. Let $\vec{p}(t) = (p_1(t), \ldots, p_N(t))^{\mathrm{T}}$ where $p_j(t) = \Pr[S(t) = j]$. $\vec{p}(t)$ is governed by the master equation

$$\frac{\mathrm{d}\vec{p}(t)}{\mathrm{d}t} = \mathbf{K}(x) \cdot \vec{p}(t),$$

where $\mathbf{K}(x) = \{k_{ij}(x)\}$ is the transition matrix. For $j \neq i$, $k_{ij}(x) = k_{j\to i}(x)$ and the diagonal elements are defined as: $k_{jj}(x) = -\sum_{i\neq j} k_{j\to i}(x)$. Thus, the transition matrix always satisfies $\sum_i k_{ij}(x) = 0$. In molecular motors, the chemical reaction is generally coordinated by the motor position. Correspondingly in the mathematical framework, the transition rates are functions of the motor position. The stochastic evolution (mechanical motion and chemical transition) of a motor system is governed by Langevin equation (4) coupled with discrete Markov process (5). The mechanical motion and the chemical transition in a two-state motor system is illustrated in Fig. 1.

Alternatively, we can follow the time evolution of probability density. Let us consider an ensemble of motors, each evolving in time independently and stochastically according to Eqs. (4) and (5). Let $\rho_s(x,t)$ be the probability density that the motor is at position $x$ and in occupancy state $S$ at time $t$. The time evolution of $\rho_s(x,t)$ is governed by the Fokker–Planck equation corresponding to (4) and (5) (Risken, 1989; Gardiner, 1985):

$$\frac{\partial \rho_i}{\partial t} = D\frac{\partial}{\partial x}\left(\underbrace{\frac{-f + \phi_i'(x)}{k_B T}\rho_i}_{\text{Convection}} + \underbrace{\frac{\partial \rho_i}{\partial x}}_{\text{Diffusion}}\right)$$
$$+ \underbrace{\sum_{j=1}^{N} k_{ij}(x)\rho_j}_{\text{Change of occupancy}}, \quad i = 1, 2, \ldots, N. \quad (6)$$

Eq. (6) is a mathematical framework for modeling the continuous motion of molecular motors (Prost et al., 1994; Peskin et al., 1994; Doering et al., 1994; Reimann, 2002; Dimroth et al., 1999; Wang and Oster, 2002; Oster and Wang, 2003). In Eq. (6), $\phi_{\mathrm{S}}(x)$'s are periodic functions with period equal to $L$ or a multiple of $L$, where $L$ is one motor step. In the subsequent sections, we are going to use this mathematical framework to study the motor potential profile and to analyze the time series of motor positions.

## 3. Motor potential profile

In many single molecule experiments, time series of motor positions were observed and recorded (Visscher et al., 1999; Yasuda et al., 1998, 2001; Block et al., 2003). In the past, these measured time series of motor positions have not been fully utilized to yield all possible information about motor mechanism. Usually only a value of average velocity and sometimes a value of randomness parameter were extracted from each time series (Visscher et al., 1999; Samuel and Berg, 1995). The extraction of the randomness parameter from time series of motor positions is a significant addition to the average velocity in deciphering motor mechanism. In the framework of kinetic models and under certain assumptions, the reciprocal of the randomness parameter tells us the number of rate-limiting chemical transitions per motor step (Visscher et al., 1999; Samuel and Berg, 1995). We believe there is much more information about motor mechanism still buried in these time series of motor positions. As we will see below, at least, a motor potential profile can be constructed from time series of motor positions.

At each position, the motor is driven by (1) the external load force, (2) the motor force derived from potential $\phi_s(x)$ corresponding to the current chemical occupancy state $S(t)$, and (3) the Brownian force from the surrounding fluid. The motor is also subject to the passive viscous drag. In the limit of high viscous friction (negligible inertia), three active forces (1), (2) and (3) are balanced by the passive viscous drag:

$$\underbrace{\zeta\frac{\mathrm{d}x}{\mathrm{d}t}}_{\substack{\text{Viscous}\\\text{drag}}} = \underbrace{f}_{\substack{\text{Load}\\\text{force}}}\ \underbrace{-\phi_S'(x)}_{\substack{\text{Motor}\\\text{force}}} + \underbrace{\sqrt{2k_B T\zeta}\,\frac{\mathrm{d}W(t)}{\mathrm{d}t}}_{\substack{\text{Brownian}\\\text{force}}}. \quad (7)$$

The external load force is usually a known constant independent of time, position and chemical state of the motor. The Brownian force is stochastic. The motor force, $\phi_s'(x)$, is also stochastic. Let us consider the sum of the three active driving forces on the right side of (7). In principle, the total active force on the motor can be measured/calculated by differentiating the time path of motor position. Then we can average the total active force at each motor position over long time regardless of chemical states. In this study, we focus on using a single potential to characterize the motor motion governed by a set of $N$ potentials. The proposed algorithm is to recover this potential from the data. This does not require resolving different chemical states from the data. If chemical states can be resolved completely or partially from the data, then more information about the set of $N$ potentials may be extracted. In averaging the total active force, the stochastic Brownian force disappears. The result is the motor force averaged over chemical states as a function of motor position, which is periodic and is called the motor force profile. The integral of the motor force profile is a tilted periodic function, and is called the motor potential profile.

For recovering the potential profile, there are practical difficulties with the approach of differentiating and averaging. For molecular motors, the stochastic instantaneous velocity caused by the Brownian force is usually larger than the average velocity. Consider the situation where a spherical bead with radius $r$, drag coefficient $\zeta$ and mass $m$, is driven by a constant force $F$. The deterministic velocity caused by the constant force $F$ is $\langle V \rangle = F/\zeta$. In addition to the constant driving force $F$, the bead is also been bombarded by the surrounding water molecules. At equilibrium, the one-dimensional root-mean-square velocity of the bead is given by the equi-partition of energy:

$$\sqrt{\mathrm{var}[V]} = \sqrt{\frac{k_B T}{m}}.$$

Using the fact that $\zeta$ is proportional to $r$ and $m$ is proportional to $r^3$, we have

$$\frac{\sqrt{\mathrm{var}[V]}}{\langle V \rangle} = \frac{\zeta}{F}\sqrt{\frac{k_B T}{m}} \sim \frac{1}{\sqrt{r}}.$$

Thus, for small beads, the fluctuations in the velocity are larger than the average velocity. Specifically, for a latex bead of 1 μm in diameter and for $F = 1pN$, we have $\sqrt{\mathrm{var}[V]}/\langle V \rangle \approx 26.7$. It takes about 700–800 data points *per motor position* to average out velocity fluctuations of this magnitude. This difficulty is compounded by other difficulties listed below:

- In the above, only the fluctuations caused by the Brownian force are counted. For molecular motors, there are additional fluctuations caused by the stochastic chemical reaction.
- In the time series, motor positions are far from being uniformly distributed. In the numerical example below, 5000 motor positions from a time series are divided into 20 bins. Even with an adaptive spatial discretization, the bin of the largest geometric size (used to accommodate the lowest density region) contains only 50 motor positions.
- Differentiating the time path of motor position is sensitive to small errors in position measurements.

In this study, our biggest advantage is that we have the mathematical framework for the underlying physical process. To overcome the difficulties listed above, we will first define of the motor potential profile in the mathematical framework of Fokker–Planck equation (6). Then we express the motor potential profile in terms of the steady-state probability flux and the steady-state probability density.

Consider the steady state of Eq. (6). Summing over $i$, we have

$$0 = D\frac{\partial}{\partial x}\left(\frac{-f + \psi'(x)}{k_B T}\rho + \frac{\partial \rho}{\partial x}\right), \tag{8}$$

where $\rho(x) = \sum_{i=1}^{N} \rho_i(x)$ is the steady-state probability density regardless of the chemical state. In the above we have used the property $\sum_{i=1}^{N} k_{ij}(x) = 0$ so that the chemical reaction terms disappear in the sum. The motor potential profile, $\psi(x)$, is defined as

$$\psi'(x) = \frac{1}{\rho(x)}\sum_{i=1}^{N} \phi_i'(x)\rho_i(x). \tag{9}$$

In Prost et al. (1994), an effective potential for a two-state model was considered in a similar way. In Eq. (8), the steady-state probability density $\rho(x)$ behaves as if the motor were driven by a single potential $\psi(x)$. In this sense, the motor potential profile represents the overall effect of the chemical reaction on the motor motion. Because $\phi_i(x)$'s are periodic with period $L$, $\psi'(x)$ is periodic and $\psi(x)$ is a tilted periodic function. $\psi(x)$ can be written as $\psi(x) = \phi(x) - \Delta\psi.x/L$, where $\phi(x)$ is a periodic function with period $L$ and $\Delta\psi = \psi(0) - \psi(L)$ can be viewed as the potential energy made available in the chemical reaction per displacement $L$ for driving the motor motion. Eq. (9) is a mathematical definition. It does not provide us a direct way of calculating the potential profile $\psi(x)$ since, in practice, both $\phi_s(x)$ and $\rho_s(x)$ are unknown. To make connection to experimental data, we rewrite Eq. (8) as

$$D\left(\frac{-f + \psi'(x)}{k_B T}\rho + \frac{\partial \rho}{\partial x}\right) = -J, \tag{10}$$

where $J$ is the steady-state probability flux, a constant independent of motor position $x$. Dividing (10) by $\rho(x)$ and integrating with respect to $x$ yields

$$\frac{\psi(x)}{k_B T} = \frac{f \cdot x}{k_B T} - \log[\rho(x)] - \frac{J}{D}\int_0^x \frac{1}{\rho(s)}\,\mathrm{d}s. \tag{11}$$

Here, we have dropped the integration constant because the potential is defined up to an additive constant. In Eq. (11), the external load force $f$ is known and the probability flux is $J = \langle v \rangle / L$, where $\langle v \rangle$ is the average velocity of the motor, which can be calculated reliably from data. Therefore, to construct the motor potential profile $\psi(x)$, we only need to extract the steady-state probability density $\rho(x)$ from the time series of motor positions. Mathematically and numerically, recovering $\rho(x)$ from the time series is certainly more viable and reliable than differentiating the time path to calculate the stochastic force.

## 4. Estimation procedure

In the previous section, we reduced the problem of constructing motor potential profile to that of recovering steady-state probability density from the time series of motor positions. In this section, we describe a robust method for doing that job.

Consider a time series of $M$ motor positions: $\{x_1, x_2, \ldots, x_M\}$. The average velocity and the probability flux are estimated as

$$\langle v \rangle \approx \frac{x_M - x_1}{t_M - t_1}, \quad J = \frac{\langle v \rangle}{L}.$$

We shift each motor position $x_j$ by an integer multiple of $L$ to make it fall in the interval $[0, L)$:

$$y_j = x_j - k_j L \in [0, L), \quad k_j = \text{integer}, \ j = 1, 2, \ldots, M.$$

A straightforward way of recovering probability density is the histogram method. Let us examine the error of the histogram method in constructing $\psi(x)$, and based on the findings we are going to propose a robust approach. Suppose the estimated probability density is $\rho(x) + \Delta\rho(x)$ where $\rho(x)$ is the exact probability density and $\Delta\rho(x)$ the absolute error and $\Delta\rho(x)/\rho(x)$ the relative error in the estimated probability density. In calculating $\psi(x)$ from $\rho(x)$ using Eq. (11), part of the corresponding change in $\psi(x)$ is $\log[\rho(x) + \Delta\rho(x)] - \log[\rho(x)]$. Using the Taylor expansion of log function, we obtain

$$\log[\rho(x) + \Delta\rho(x)] - \log[\rho(x)] = \log\left[1 + \frac{\Delta\rho(x)}{\rho(x)}\right] \approx \frac{\Delta\rho(x)}{\rho(x)}. \tag{12}$$

Eq. (12) shows that the relative error in the estimated $\rho(x)$ becomes the absolute error in the estimated $\psi(x)$. The relative error in the estimated $\rho(x)$ comes from two sources: the spatial discretization error (affected by the bin size we use) and the statistical error (affected by how many data points we have in a bin). Here we focus on the statistical error. In the histogram method, the relative statistical error in the estimated $\rho(x)$ is approximately given by (see Appendix A for details)

$$\frac{\Delta\rho(x)}{\rho(x)} \sim \frac{1}{\sqrt{n_x}}, \tag{13}$$

where $n_x$ is the average number of data points ($y_j$'s) in the bin around $x$. If we use bins of equal size, then $n_x$ is approximately proportional to $\rho(x)$. As a result, the relative statistical error in the estimated $\rho(x)$ is large in regions where $\rho(x)$ is small. To keep the relative error uniformly small, we should use bins of variable sizes according to the data so that each bin contains about the same number of data points. The probability density function estimated in the histogram method is a discontinuous step function. It is not clear how to interpolate this step function to make it continuous while keeping it positive and integrating to one. Alternatively, one may assume that $\rho(x)$ is continuous and piecewise linear, and use the maximum likelihood method to estimate $\rho(x)$. Although the $\rho(x)$ estimated this way is continuous, in the region where the true value of $\rho(x)$ is very small, the estimated $\rho(x)$ may be negative, which will be disaster for constructing $\psi(x)$ using (11). Also given that the true value of $\rho(x)$ may differ by orders of magnitude at different locations, linear interpolation does not offer the best way for representing $\rho(x)$. To motivate a more robust and more efficient way of representing $\rho(x)$, let us consider the Boltzmann distribution corresponding to potential $\phi(x)$:

$$\rho(x) \propto \exp\left(\frac{-\phi(x)}{k_B T}\right).$$

Potential $\phi(x)$ is not constrained to be positive so it is reasonable to approximate it using a continuous piecewise linear function. The corresponding $\rho(x)$ is a continuous piecewise exponential function, which guarantees the positivity. Also a piecewise exponential probability density makes the maximum likelihood formulation easy to solve because multiplying probability densities corresponds to adding exponents.

Based on the analysis above, we propose the method described below for estimating $\rho(x)$ and reconstructing $\psi(x)$:

*Step* 1: We divide the interval $[0, L)$ into cells according to the data $\{y_1, y_2, \ldots, y_M\}$:

$$0 = a_0 < a_1 < a_2 < \cdots < a_{m-1} < a_m = L.$$

To keep the relative error uniformly small, we require that each cell $[a_{j-1}, a_j)$ contains at least $m_2$ data points. For computational efficiency, we also require that the size of each cell $(a_j - a_{j-1})$ is not smaller than a prescribed size $\Delta x$. In this way, we can avoid having lots of small cells concentrated in the region where $\rho(x)$ is large. Specifically, we start with $a_0 = 0$. Once we have $a_j$, we select $a_{j+1}$ as

$$a_{j+1} = \min\{a | a \geqslant a_j + \Delta x \text{ and } [a_j, a) \text{ contains at least } m_2 \text{ data points }\}.$$

Note that $m$ is not prescribed. Rather, $m$ is determined by $m_2$ and $\Delta x$. We should choose $m_2$ and $\Delta x$ to make $m$ (the number of cells) much smaller than $M$ (the number of data points).

*Step* 2: For maintaining the positivity and for accurate spatial approximation, we assume that $\rho(x)$ is exponential in each cell and is continuous at cell boundaries. Mathematically, we assume $\rho(x)$ is completely specified by $b_0, b_1, \ldots, b_m$, and has the form

$$\rho(x) = \exp\left[b_{j-1} + \frac{b_j - b_{j-1}}{a_j - a_{j-1}}(x - a_{j-1})\right],$$
$$x \in [a_{j-1}, a_j],\ \ j = 1, 2, \ldots, m, \tag{14}$$

where $b_j$ represents $\log[\rho(a_j)]$. Because $\rho(x)$ is probability density, $b_j$'s satisfy

$$\sum_{j=1}^{m} \int_{a_{j-1}}^{a_j} \exp\left[b_{j-1} + \frac{b_j - b_{j-1}}{a_j - a_{j-1}}(x - a_{j-1})\right] \mathrm{d}x = 1,$$

which leads to

$$S(b_1, b_2, \ldots, b_m) \equiv \sum_{j=1}^{m} \frac{a_j - a_{j-1}}{b_j - b_{j-1}} [\exp(b_j) - \exp(b_{j-1})] = 1. \tag{15}$$

Since $\rho(x)$ is periodic, we have $b_0 = b_m$. So we only need to determine $(b_1, b_2, \ldots, b_m)$.

We use the maximum likelihood method to estimate $\rho(x)$. We collect data points falling in each cell $[a_{j-1}, a_j)$ into a group. We have $m$ groups of data points:

Cell 1: $y_{1,1}, y_{1,2}, \ldots, y_{1,M_1} \in [a_0, a_1)$

Cell 2: $y_{2,1}, y_{2,2}, \ldots, y_{2,M_2} \in [a_1, a_2)$

$\vdots$

Cell $m$ : $y_{m,1}, y_{m,2}, \ldots, y_{m,M_m} \in [a_{m-1}, a_m)$.

The likelihood of observing the data $\{y_1, y_2, \ldots, y_M\}$ is

$$\begin{aligned} L(b_1, b_2, \ldots, b_m) &\equiv \prod_{i=1}^{M} \rho(y_i | b_1, b_2, \ldots, b_m) \\ &= \prod_{j=1}^{m} \prod_{i=1}^{M_j} \exp\left[b_{j-1} + \frac{b_j - b_{j-1}}{a_j - a_{j-1}}(y_{j,i} - a_{j-1})\right] \\ &= \prod_{j=1}^{m} \exp[(M_j - q_j) b_{j-1} + q_j b_j] \\ &= \exp\left[M \sum_{j=1}^{m} c_j b_j\right], \end{aligned} \tag{16}$$

where

$$q_j = \frac{1}{a_j - a_{j-1}} \sum_{i=1}^{M_j} (y_{j,i} - a_{j-1}), \quad j = 1, 2, \ldots, m,$$

$$c_j = \frac{1}{M}(M_{j+1} - q_{j+1} + q_j) > 0, \quad j = 1, 2, \ldots, m-1,$$

$$c_m = \frac{1}{M}(M_1 - q_1 + q_m) > 0.$$

In our method, $\{c_1, c_2, \ldots, c_m\}$ along with the numerical grid $\{a_0, a_1, \ldots, a_m\}$ is a sufficient statistics of the data. $\{c_1, c_2, \ldots, c_m\}$ satisfies

$$\sum_{j}^{m} c_j = \frac{1}{M} \sum_{j}^{m} M_j = 1.$$

The use of likelihood function here needs more explanation and is debatable. Strictly speaking, (16) is rigorously valid only when $\{y_1, y_2, \ldots, y_M\}$ are independent. The motor positions *at consecutive times* from a time series certainly are not independent. For example, $y_{j+1}$ is far from being independent of $y_j$. However, notice that the function given in (16) depends only on the sufficient statistics $\{c_1, c_2, \ldots, c_m\}$ of the data. In particular, the function given in (16) does not depend on the order of the sequence $\{y_1, y_2, \ldots, y_M\}$. That is, if we *permute the sequence* $\{y_1, y_2, \ldots, y_M\}$, the function given in (16) is not affected at all. So instead of applying the function given in (16) to the data sequence $\{y_1, y_2, \ldots, y_M\}$ directly, we first do a random permutation on the data sequence $\{y_1, y_2, \ldots, y_M\}$ to obtain a set of randomly permuted observations $\{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_M\}$. For large $M$, the randomly permuted observations $\{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_M\}$ are approximately independent. Thus, it is reasonable to apply the likelihood function given in (16) to the randomly permuted observations $\{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_M\}$. Another supporting argument for using the likelihood method in this situation is that the histogram method, which does not require independence, can be formulated in the form of the likelihood method.

The logarithm of the likelihood function is

$$F(b_1, b_2, \ldots, b_m) \equiv \frac{1}{M} \log[L(b_1, b_2, \ldots, b_m)] = \sum_{j=1}^{m} c_j b_j.$$

*Step* 3: $(b_1, b_2, \ldots, b_m)$ is determined by maximizing the logarithm of the likelihood function (Harris and Stocker, 1988; Hoel, 1962; Fisher, 1922)

$$\underset{S(b_1, b_2, \ldots, b_m) = 1}{\operatorname{argmax}} \ F(b_1, b_2, \ldots, b_m). \tag{17}$$

In (17), the objective function is linear while the constraint is non-linear. We switch the roles of $S(b_1, b_2, \ldots, b_m)$ and $F(b_1, b_2, \ldots, b_m)$, and consider the minimization problem

$$\underset{F(b_1, b_2, \ldots, b_m) = 0}{\operatorname{argmin}} \ S(b_1, b_2, \ldots, b_m). \tag{18}$$

We conclude that if $(\beta_1, \beta_2, \ldots, \beta_m)$ is the solution of problem (18), then

$$(\beta_1, \beta_2, \ldots, \beta_m) + \delta \quad \text{where } \delta = -\log[S(\beta_1, \beta_2, \ldots, \beta_m)] > 0 \tag{19}$$

is the solution of problem (17) (see Appendix B for derivation). So we only need to solve problem (18). Since $F(b_1, b_2, \ldots, b_m)$ is linear, we rewrite problem (18) as an unconstrained minimization problem

$$\underset{(b_1, b_2, \ldots, b_{m-1})}{\operatorname{argmin}} \ S\left(b_1, \ldots, b_{m-1}, \frac{-1}{c_m} \sum_{j=1}^{m-1} c_j b_j\right). \tag{20}$$

There are many good numerical methods for solving unconstrained non-quadratic minimization problems. Here we use the non-linear conjugate gradient method (Fletcher and Reeves, 1964; Polak and Ribiere, 1969). Once the solution, $(\beta_1, \ldots, \beta_{m-1})$, of problem (20) is computed, we set

$$\beta_m = \frac{-1}{c_m} \sum_{j=1}^{m-1} c_j \beta_j, \quad \delta = -\log[S(\beta_1, \beta_2, \ldots, \beta_m)],$$

$$(b_1, b_2, \ldots, b_m) = (\beta_1, \beta_2, \ldots, \beta_m) + \delta.$$

Then $\mathbf{b} = (b_1, b_2, \ldots, b_m)$ is the solution of non-linearly constrained maximization problem (17). The corresponding probability density is given by

$$\rho(x|\mathbf{b}) = \exp\left[b_{j-1} + \frac{b_j - b_{j-1}}{a_j - a_{j-1}}(x - a_{j-1})\right],$$
$$x \in [a_{j-1}, a_j], \; j = 1, 2, \ldots, m.$$

Let $\boldsymbol{\psi}(\mathbf{b}) = \{\psi(a_j|\mathbf{b}), \; j = 0, 1, \ldots, m\}$ be the corresponding potential profile on the numerical grid $\{a_0, a_1, \ldots, a_m\}$. $\boldsymbol{\psi}(\mathbf{b})$ is a vector given by

$$\begin{aligned}\frac{\psi(a_j|\mathbf{b})}{k_B T} &= \frac{f \cdot a_j}{k_B T} - \log[\rho(a_j|\mathbf{b})] - \frac{J}{D} \int_0^{a_j} \frac{1}{\rho(s|\mathbf{b})} \mathrm{d}s \\ &= \frac{f \cdot a_j}{k_B T} - b_j - \frac{J}{D} \sum_{l=1}^{j} \frac{a_l - a_{l-1}}{(-b_l) - (-b_{l-1})} \\ &\quad \times [\exp(-b_l) - \exp(-b_{l-1})]. \end{aligned} \tag{21}$$

*Step* 4: We now estimate the statistical error in $\mathbf{b} = (b_1, b_2, \ldots, b_m)$ and the statistical error in $\boldsymbol{\psi}(\mathbf{b}) = \{\psi(a_j|\mathbf{b}), \; j = 0, 1, \ldots, m\}$. We first introduce several shorthand notations.

Let $\mathbf{y} = \{y_1, y_2, \ldots, y_M\}$ denote the given data, shifted into one period.

Let $\mathbf{b}(\mathbf{y}) = (b_1(\mathbf{y}), b_2(\mathbf{y}), \ldots, b_m(\mathbf{y}))$ denote the solution of (17) using data $\mathbf{y}.\mathbf{b}(\mathbf{y})$ is a vector.

Let $\rho(x|\mathbf{b}(\mathbf{y}))$ denote the probability density estimated using data $\mathbf{y}.\rho(x|\mathbf{b}(\mathbf{y}))$ is a function.

Let $\boldsymbol{\psi}(\mathbf{b}(\mathbf{y})) = \{\psi(a_j|\mathbf{b}(\mathbf{y})), \; j = 0, 1, \ldots, m\}$ denote the potential profile on the numerical grid $\{a_0, a_1, \ldots, a_m\}$, estimated using data $\mathbf{y}.\boldsymbol{\psi}(\mathbf{b}(\mathbf{y}))$ is a vector.

Let $\mathbf{d} = \{d_1, d_2, \ldots, d_M\}$ be a set of $M$ iid (independently identically distributed) random numbers in $[0, L]$, drawn according to the estimated probability density $\rho(x|\mathbf{b}(\mathbf{y}))$. In the example below, the straightforw ard accept/reject method is used in generating $\mathbf{d} = \{d_1, d_2, \ldots, d_M\}$. If the efficiency of the accept/reject method is too low, the inverse function of the cumulative distribution function can be used to map a set of uniformly distributed random numbers to generate $\mathbf{d} = \{d_1, d_2, \ldots, d_M\}$.

Using data $\mathbf{d}$, we solve (17) for $\mathbf{b}(\mathbf{d})$. Note that if we draw an infinite set of iid random numbers according to $\rho(x|\mathbf{b}(\mathbf{y}))$ and use this infinite data set in solving (17), then $\mathbf{b}(\mathbf{y})$ will be recovered exactly. The statistical error in $\mathbf{b}(\mathbf{y})$, caused by the finite sample size, is estimated as the standard deviation of $\mathbf{b}(\mathbf{d})$:

$$\mathrm{Error}[b_j(\mathbf{y})] \sim \mathrm{std}[b_j(\mathbf{d})]_{\mathbf{d}}, \quad j = 1, 2, \ldots, m,$$

where $\mathbf{d}$ is treated as a random variable and the average is taken with respect to $\mathbf{d}$.

Similarly, the statistical error in $\boldsymbol{\psi}(\mathbf{b}(\mathbf{y}))$ is estimated as the standard deviation of $\boldsymbol{\psi}(\mathbf{b}(\mathbf{d}))$

$$\mathrm{Error}[\psi(a_j|\mathbf{b}(\mathbf{y}))] \sim \mathrm{std}[\psi(a_j|\mathbf{b}(\mathbf{d}))]_{\mathbf{d}}, \quad j = 1, 2, \ldots, m.$$

More specifically, in the numerical implementation, we draw $n$ independent sets of random numbers, each set consisting of $M$ iid random numbers drawn according to $\rho(x|\mathbf{b}(\mathbf{y}))$:

$$\mathbf{d}^{(1)} = \{d_1^{(1)}, d_2^{(1)}, \ldots, d_M^{(1)}\},$$

$$\mathbf{d}^{(2)} = \{d_1^{(2)}, d_2^{(2)}, \ldots, d_M^{(2)}\},$$

$$\vdots$$

$$\mathbf{d}^{(n)} = \{d_1^{(n)}, d_2^{(n)}, \ldots, d_M^{(n)}\}$$

with each data set $\mathbf{d}^{(i)}$, $i = 1, 2, \ldots, n$, we solve (17) for $\mathbf{b}(\mathbf{d}^{(i)})$. In solving (17), we keep the numerical grid $\{a_0, a_1, \ldots, a_m\}$ unchanged. That is, we use the numerical grid calculated from data $\mathbf{y}$; we do not recalculate the numerical grid for data $\mathbf{d}^{(i)}$. The statistical error in $b_j(\mathbf{y})$ is estimated as

$$\langle b_j(\mathbf{d}) \rangle_{\mathbf{d}} \approx \frac{1}{n} \sum_{i=1}^{n} b_j(\mathbf{d}^{(i)}), \quad j = 1, 2, \ldots, m,$$

$$\mathrm{std}[b_j(\mathbf{d})]_{\mathbf{d}} \approx \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (b_j(\mathbf{d}^{(i)}) - \langle b_j(\mathbf{d}) \rangle_{\mathbf{d}})^2},$$
$$j = 1, 2, \ldots, m. \tag{22}$$

The statistical error in $\psi(a_j|\mathbf{b}(\mathbf{y}))$ is estimated as

$$\langle \psi(a_j|\mathbf{b}(\mathbf{d})) \rangle_{\mathbf{d}} \approx \frac{1}{n} \sum_{i=1}^{n} \psi(a_j|\mathbf{b}(\mathbf{d}^{(i)})), \quad j = 1, 2, \ldots, m,$$

$$\mathrm{std}[\psi(a_j|\mathbf{b}(\mathbf{d}))]_{\mathbf{d}}$$
$$\approx \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (\psi(a_j|\mathbf{b}(\mathbf{d}^{(i)})) - \langle \psi(a_j|\mathbf{b}(\mathbf{d})) \rangle_{\mathbf{d}})^2},$$
$$j = 1, 2, \ldots, m. \tag{23}$$

Now let us summarize the method proposed above:

(1) Calculate the numerical grid $\{a_0, a_1, \ldots, a_m\}$ using the given data $\mathbf{y}$.
(2) Calculate the sufficient statistics $\{c_1, c_2, \ldots, c_m\}$ using data $\mathbf{y}$ and using the numerical grid $\{a_0, a_1, \ldots, a_m\}$ from 1.

(3) Solve (17) for $\mathbf{b}(\mathbf{y}) = (b_1, b_2, \ldots, b_m)$ using $\{c_1, c_2, \ldots, c_m\}$ calculated from data $\mathbf{y}$ and the numerical grid $\{a_0, a_1, \ldots, a_m\}$ from 1.

(4) Estimate the statistical errors in $\mathbf{b}(\mathbf{y})$ and $\boldsymbol{\psi}(\mathbf{b}(\mathbf{y}))$ by carrying out steps below:

  (a) Draw $n$ artificial data sets $\{\mathbf{d}^{(i)}\}$ according to $\rho(x|\mathbf{b}(\mathbf{y}))$.

  (b) Repeat 2 and 3 above with data $\mathbf{y}$ replaced by data $\mathbf{d}^{(i)}$ to calculate $\mathbf{b}(\mathbf{d}^{(i)})$.

  (c) Calculate the mean and standard deviation of $\{\mathbf{b}(\mathbf{d}^{(i)})\}$.

  (d) Calculate the mean and standard deviation of $\{\boldsymbol{\psi}(\mathbf{b}(\mathbf{d}^{(i)}))\}$.

## 5. Numerical results

To test the method described in the previous section, we consider a hypothetical motor system switching mathematically between two periodic potentials. The two potentials are periodic with period $2L$, and are given by

$$\phi_1(x) = A\begin{cases} \left(\dfrac{4}{5L}x - 1\right)^2, & 0 < x < \dfrac{5}{4}L, \\ 0, & \dfrac{5}{4}L < x < \dfrac{6}{4}L, \\ \dfrac{2}{L}\left(x - \dfrac{6}{4}L\right), & \dfrac{6}{4}L < x < 2L, \end{cases}$$

$$\phi_2(x) = \phi_1(x - L) - \Delta G.$$

From the point of view of energetics, the motor system is actually moving along an infinite sequence of sets of potentials $(\phi_{0,1}(x), \phi_{0,2}(x))$, $(\phi_{1,1}(x), \phi_{1,2}(x))$, $(\phi_{2,1}(x), \phi_{2,2}(x)), \ldots, (\phi_{n,1}(x), \phi_{n,2}(x)), \ldots$. Each set of potentials corresponds to a reaction cycle. For the $n$th reaction cycle, the two potentials are related to $\phi_1(x)$ and $\phi_2(x)$ by

$$\phi_{n,1}(x) = \phi_1(x) - 2n\Delta G,$$

$$\phi_{n,2}(x) = \phi_2(x) - 2n\Delta G,$$

where $-2\Delta G$ is the free energy change of one reaction cycle. As shown in Fig. 2, after the chemical reaction switches the motor system onto $\phi_{0,1}(x)$ (state 1 of cycle 0), the motor system is driven by potential $\phi_{0,1}(x)$ towards its minimum energy position. Near the minimum energy position of $\phi_{0,1}(x)$, the reaction switches the motor onto $\phi_{0,2}(x)$ (state 2 of cycle 0), and $\phi_{0,2}(x)$ derives the motor towards its minimum energy position. Near the minimum energy position of $\phi_{0,2}(x)$, the reaction switches the motor onto $\phi_{1,1}(x)$ (state 1 of cycle 1), and the process is repeated (the motor proceeds from $\phi_{1,1}(x)$ to $\phi_{2,1}(x)$, not shown in the figure).

From the point of view of energetics, it is important to distinguish potentials of different cycles. For example, the switch from $\phi_{0,2}(x)$ to $\phi_{1,1}(x)$ is a forward transition in the positive direction of reaction, which carries the motor to the next reaction cycle, while the switch from $\phi_{0,2}(x)$ to $\phi_{0,1}(x)$ is a backward transition. In the hypothetical motor system shown in Fig. 2, we assume that chemical reaction is well coordinated by the motor position. Specifically, we assume the chemical transition rates are periodic with period $2L$, and are given by

$$k_{(0,1)\to(0,2)}(x) = k_0\begin{cases} \exp\left(\dfrac{\phi_1(x)}{k_BT}\right), & \dfrac{11}{8}L < x < \frac{13}{8}L, \\ 0, & 0 < x < \dfrac{11}{8}L \text{ or} \\ & \dfrac{13}{8}L < x < 2L, \end{cases}$$

$$k_{(0,2)\to(0,1)}(x) = k_{(0,1)\to(0,2)}(x)\exp\left(\frac{\phi_2(x) - \phi_1(x)}{k_BT}\right),$$

$$k_{(0,2)\to(1,1)}(x) = k_{(0,1)\to(0,2)}(x - L),$$

$$k_{(1,1)\to(0,2)}(x) = k_{(0,2)\to(0,1)}(x - L).$$

Mathematically, in the Fokker–Planck equation, we treat the collection $\{\phi_{n,1}, -\infty < n < \infty\}$ as one composite state (occupancy state $S_1$) and treat the collection $\{\phi_{n,2}, -\infty < n < \infty\}$ as another composite state (occupancy state $S_2$). Note that if we focus on the occupancy of the catalytic site and disregard the free energy of system, $\phi_{0,1}$ and $\phi_{1,1}$ are not distinguishable. The two occupancy states and the transitions between them are illustrated in Fig. 3.

As shown in Fig. 3, the transition rate from occupancy state $S_2$ to occupancy state $S_1$ includes both the transition from state 2 to state 1 of the same cycle and the transition from state 2 of the current cycle to state 1 of the next cycle.

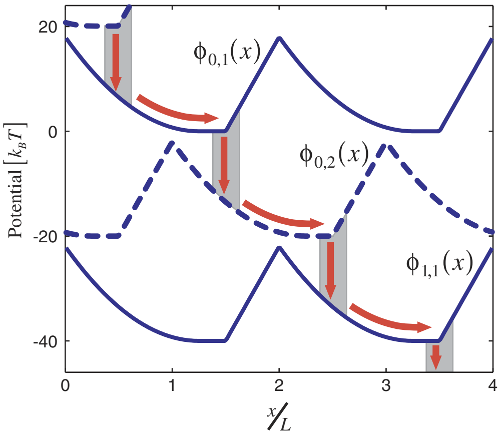

Fig. 2. A hypothetical motor system. From the point of view of energetics, the motor system moves along an infinite sequence of sets of potentials, each set consisting of two potentials ($\phi_{n,1}(x)$ and $\phi_{n,2}(x)$) and corresponding to one reaction cycle. The chemical transition is well coordinated by motor position, and is only allowed to occur in the shaded regions.

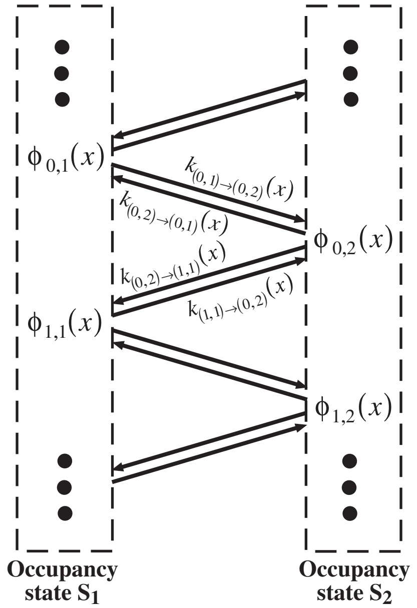

Fig. 3. Occupancy states used in the Fokker–Planck equation. From the point of view of energetics, as the chemical reaction proceeds, the motor system in Fig. 2 goes through a sequence of sets of potentials ($\phi_{n,1}(x)$ and $\phi_{n,2}(x)$) with free energy decreasing. If we focus on the occupancy of catalytic site and disregard the free energy of system, we can treat $\{\phi_{n,1}, -\infty<n<\infty\}$ and $\{\phi_{n,2}, -\infty<n<\infty\}$ as two composite states so that the motor system switches mathematically between two composite states in the Fokker–Planck equation.

In the Fokker–Planck equation, the transition rates between $S_1$ and $S_2$ are given by

$$k_{2\to1}(x) = k_{(0,2)\to(1,1)}(x) + k_{(0,2)\to(0,1)}(x),$$

$$k_{1\to2}(x) = k_{(0,1)\to(0,2)}(x) + k_{(1,1)\to(0,2)}(x).$$

In the simulations, we use parameters listed below:

$$A = 20k_BT, \quad \Delta G = 20k_BT, \quad L = 8\,\text{nm},$$
$$D = 10^4\,\text{nm}^2/\text{s}, \quad k_0 = 10^3/\text{s}.$$

We first integrate in time numerically the Langevin equation (4) and the discrete Markov process (5) using potentials, transitions rates and parameters given above. In the Langevin simulations, we record a time series of 5000 motor positions over a time period of 4 s, which corresponds to a sampling rate of 1.25 kHz in experiments. We treat this simulated time series of 5000 motor positions as data and use it to test the method for extracting motor potential profile. In the process of extracting motor potential profile, we use parameters: $m_2 = 50$ (minimum number of data points in each cell) and $\Delta x = L/20 = 0.4$ (minimum size of cells) unless otherwise noted. Fig. 4 shows the exact motor potential profile calculated using the definition and the motor potential profile estimated from the simulated time series of 5000 motor positions. The analytic solution of Fokker–Planck equation (6) is not known. Here the "exact solution" means a numerical solution of (6) using the robust numerical method developed in Wang et al. (2003) with a fine numerical grid ($N = 2048$ points per period). In Fig. 4, two periods of the motor potential profile are plotted. If only one period is plotted, then the power stroke will be split into two parts: one near the right end of the period and the other near the left end. The exact motor potential profile captures the essential features of the motor mechanism for the hypothetical motor system. Between energy barriers, a powerstroke drives the motor. As the motor moves forward, the magnitude of the powerstroke decreases and the powerstroke is followed by a flat step in free energy. The flat step does not provide an active driving force and the motor depends on Brownian diffusion to get over the flat step. The energy barrier after the flat step is caused by the fact that the transitions between potentials are one of the rate-limiting steps for the motor system. While the motor system is waiting on $\phi_{0,1}(x)$ for transition to $\phi_{0,2}(x)$, for example, it diffuses around and feels the energy barrier on $\phi_{0,1}(x)$ (see Fig. 2). As shown in Fig. 4, it is clear that the estimated motor potential profile matches very well the exact one and captures all the essential features of the motor mechanism described above. In particular, the powerstroke of decreasing magnitude and the flat step with no active driving force are well captured in the estimated potential profile.

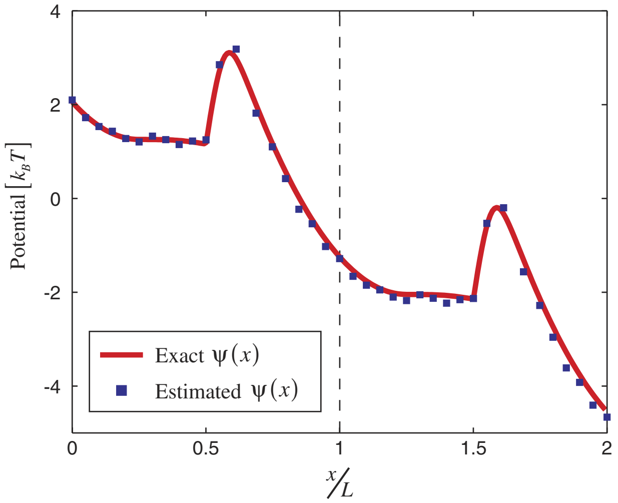

Fig. 4. The exact potential profile and the potential profile estimated from a time series of 5000 motor positions. The potential is continuous and piecewise smooth. Values of the estimated potential at numerical grid points are shown. The estimated potential profile matches the exact one very well.

Fig. 5 shows the exact $\log(\rho(x))$, the estimated $\log(\rho(x))$ and the associated statistical error due to finite sample size. Again, the "exact solution" means a very accurate numerical solution as described above. The error bars shown represent interval [mean −std, mean + std]. The error bars for $\log(\rho(x))$ are calculated using (22). If the underlying random variable is Gaussian, then the probability of falling in interval [mean −std, mean + std] is 68.27%, and the probability of falling in interval [mean $-2 \times$ std, mean $+ 2 \times$ std] is 95.45%. Although the

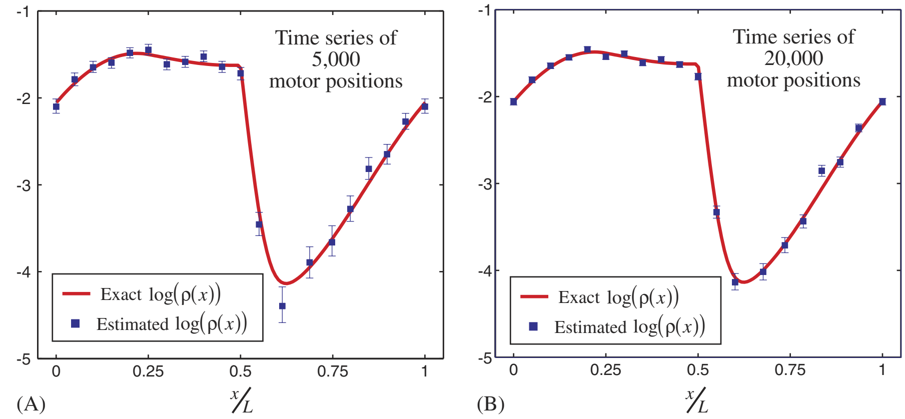

Fig. 5. The exact $\log(\rho(x))$, the estimated $\log(\rho(x))$ and the associated statistical error due to finite sample size. $\log(\rho(x))$ is continuous and piecewise linear. Values of the estimated $\log(\rho(x))$ at numerical grid points are shown. Between two adjacent numerical grid points, $\log(\rho(x))$ is linear. To estimate the statistical error, $n = 1001$ data sets are generated, each data set consisting of 5000 points drawn according to $\rho(x)$. One sample for the estimated $\log(\rho(x))$ is calculated from each data set. The error bars represent interval [mean $-$std, mean $+$ std]. The error bars for $\log(\rho(x))$ are calculated using (22). Left panel: results for a time series of 5000 motor positions. Right panel: results for a time series of 20,000 motor positions.

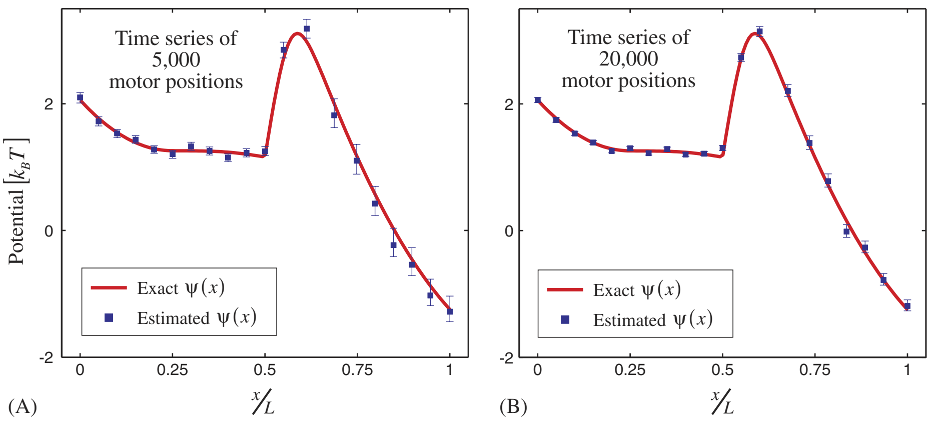

Fig. 6. The exact $\psi(x)$, the estimated $\psi(x)$ and the associated statistical error due to finite sample size. The procedure for calculating the statistical error is described in the text and in the caption of Fig. 5. $n = 1001$ sets of generated data are used in the error analysis. The error bars represent interval [mean $-$std, mean $+$ std]. The error bars for $\psi(x)$ are calculated using (23). Left panel: results for a time series of 5000 motor positions. Right panel: results for a time series of 20,000 motor positions.

estimated $\log(\rho(a_j))$ may not be exactly Gaussian, it behaves very much like a Gaussian random variable (results not shown). In Fig. 5, results are shown for a time series of 5000 motor positions (left panel) and a time series of 20,000 motor positions (right panel). It should be pointed out that these two time series are not independent of each other: the time series of 5000 motor positions is formed simply by taking one out of every four motor positions in the time series of 20,000 motor positions. So the results for these two time series may show similar bias in some regions. The general trend is clear: as the sample size increases the statistical error converges to zero and the estimated $\log(\rho(x))$ converges to the exact one. Fig. 6 shows the exact $\psi(x)$ (potential profile), the estimated $\psi(x)$ and the associated statistical error due to finite sample size. The error bars shown represent interval [mean $-$std, mean $+$ std]. The error bars for $\psi(x)$ are calculated using (23). Results for two time series are shown. It is clear that as the sample size increases the estimated potential profile converges to the exact one. As given in (21), when the external force ($f$) is zero, the estimated potential profile has two parts: $b_j$ and a cumulative sum involving $(b_0, b_1, \ldots, b_j)$. Both of these two parts contribute to the statistical error of the estimated potential profile at $a_j$. As the index $j$ increases, the cumulative sum consists of more and more terms and theassociated statistical error increases. This behavior is demonstrated in Fig. 6. This behavior is neither surprising nor unreasonable. The potential profile is only determined up to an additive constant. We have the choice of fixing the potential profile at a point. Near that point, the statistical error is small. As we get away from that point, the statistical error increases. In the mathematical formulation (11) and the implementation (21), roughly (not exactly) we fix the potential profile at the left end ($j = 0$). So it is reasonable to expect the statistical error to increase as $j$ increases.

## 6. Conclusion

In this paper, we propose a robust method for extracting the motor potential profile from time series of motor positions. The motor potential profile was proposed in Wang (2003) as a simple and practical way of characterizing the motor mechanism using one potential curve instead of using a set of $N$ potential curves. Mathematically, the motor potential profile is defined by averaging motor force over all chemical occupancy states at each motor position. The chemical reaction in a molecular motor has many occupancy states, each having a different effect on the motor motion. The overall effect of the chemical reaction on the motor motion is characterized by the motor potential profile. The potential profile reveals how the motor force changes with position in a motor step, which will lead to insights into how the chemical reaction is coupled to force generation. As the single molecule experimental technology advances, it allows us to measure motor position and force with an accuracy of sub-nanometer and sub-piconewton (Block et al., 2003; Abbondanzieri et al., 2005). But it has not yet allowed us to record motor position and chemical occupancy state at the same time. Therefore, it is still unrealistic to extract a full set of $N$ potential curves, one for each chemical occupancy state. The best advantage of studying the motor potential profile is that it does not require us to distinguish/resolve different chemical occupancy states, and thus, in principle, it can be extracted from time series of motor positions with no explicit information on the current chemical state. The proposed method is based on the mathematical formulation that expresses the motor potential profile in terms of steady-state probability flux and steady-state probability density. Furthermore, the proposed method adopts a robust numerical framework (1) to guarantee the positivity of the probability density, (2) to maintain a uniform accuracy especially in the region where the probability density is low, and (3) to estimate the statistical error due to the finite sample size. The proposed method is tested using the simulated data from a hypothetical motor system. The test results demonstrate that the proposed method is capable of reconstructing a motor potential profile with a reasonable accuracy from a time series of moderate size (5000 motor positions). In addition, the proposed method also gives the statistical errors associated with the estimated potential profile. The next step will be to apply the proposed method, in collaboration with experimentalists, to extracting motor potential profile from time series of motor positions measured in real single molecule experiments.

Recently Walton proposed to use the hidden Markov model to analyze single molecule experimental data on kinesin (Walton, 2002). Both Walton's study and our study share the same general approach of combining mathematical models describing the motor with statistical methods to process the data. These two studies also share the same goal of extracting more information about the motor from the data. The specific goals and the specific mathematical models and statistical methods used are different. In Walton's study, the kinesin is described by a kinetic model with a few states; the statistical tool is the hidden Markov model; and the goal is to extract the time series of kinesin states from the measured time series of bead positions (that is, the goal is to build a hidden Markov model filter with the measured time series of bead positions as the input and the recovered time series of kinesin states as the output). Since the goal is to extract the stochastic time series of kinesin states instead of extracting an average quantity, it requires high time resolution (that is, it requires many data points near each time instance). In Walton's study, it was found that the time resolution of the current experimental data is not high enough for extracting the intermediate states (Walton, 2002). This, however, should not be viewed as a defect of Walton's study at all. Rather, it should be viewed as a design guide for future experiments and for future experimental technologies. In our study, the motor is described by Fokker–Planck equation (or equivalently a Langevin equation); the statistical tool is the likelihood method; and the goal is to extract the motor potential profile, which is an average entity (it is a function instead of a number). Both Walton's study and our study demonstrate that when the general statistical methods are combined with the specific mathematical models describing the underlying physical process, more information may be extracted.

## Acknowledgments

The author thanks George Oster and Tim Elston for helpful discussions during the project, and thanks anonymous reviewers for their constructive suggestions and comments in revising and improving the manuscript. This work was partially supported by NSF Grant DMS-0317937.

## Appendix A. Derivation of Eq. (13)

In this appendix, we estimate the relative statistical error in the estimated probability density using the histogram method. For simplicity, we treat $\{y_1, y_2, \ldots, y_M\}$ as iid (independently identically distributed) random variables with probability density $\rho(x)$. The assumption that $\{y_1, y_2, \ldots, y_M\}$ are iid needs explanation. Apparently, $\{y_1, y_2, \ldots, y_M\}$ are not iid. For example, $y_{j+1}$ is far from being independent of $y_j$. In the histogram method, however, the result depends only on a sufficient statistics of $\{y_1, y_2, \ldots, y_M\}$, i.e. how many data points in each bin. In particular, the result does not depend on the order of the sequence $\{y_1, y_2, \ldots, y_M\}$. That is, if we *permute the sequence* $\{y_1, y_2, \ldots, y_M\}$, the result of the histogram method is not affected at all. Therefore, instead of applying the histogram method directly on the data $\{y_1, y_2, \ldots, y_M\}$, we first do a random permutation on the data sequence $\{y_1, y_2, \ldots, y_M\}$ to obtain a set of randomly permuted observations $\{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_M\}$. For large $M$, the randomly permuted observations $\{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_M\}$ are approximately iid. Then we apply the histogram method on the randomly permuted observations $\{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_M\}$. So the real assumption we make here is that the randomly permuted observations $\{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_M\}$ is iid, which is approximately true.

Let $[a, b)$ be the bin around $x$ and let $p = \int_a^b \rho(x)\,\mathrm{d}x$. If $[a, b)$ is one of many bins used in the histogram, then we have $p \ll 1$. The histogram method estimates $p$ and $\rho(x)$ as

$$P = \frac{\#\ \text{of}\{y_1, y_2, \ldots, y_M\}\ \text{in}[a, b)}{M}, \quad \rho(x) \approx \frac{P}{b-a}.$$

Let us introduce random variable $R_j$:

$$R_j = \begin{cases} 1, & y_j \in [a, b), \\ 0 & \text{otherwise}, \end{cases} \quad \langle R_j \rangle = p, \ \text{var}(R_j) = p(1-p)$$

$P$ (the estimated value of $p$) is a random variable. It can be written in terms of $R_j$ as

$$P = \frac{1}{M} \sum_{j=1}^{M} R_j, \quad \langle P \rangle = p, \ \text{var}(P) = \frac{1}{M} p(1-p).$$

The relative statistical error in the estimated $\rho(x)$ is roughly

$$\frac{|\Delta\rho(x)|}{\rho(x)} \sim \frac{\sqrt{\text{var}(P)}}{\langle P \rangle} = \sqrt{\frac{1-p}{Mp}} \approx \frac{1}{\sqrt{Mp}} \equiv \frac{1}{\sqrt{n_x}},$$

where $n_x = Mp$ is the average number of data points (out of total $M$ data points) in the bin $[a, b)$.

## Appendix B. Relation between problem (17) and (18)

Suppose $(\beta_1, \beta_2, \ldots, \beta_m)$ is the solution of (18). By definition, it satisfies

$$F(\beta_1, \beta_2, \ldots, \beta_m) = 0,$$

$$S(\beta_1, \beta_2, \ldots, \beta_m) \leqslant S[(\beta_1, \beta_2, \ldots, \beta_m) + (\Delta\beta_1, \ldots, \Delta\beta_m)] \tag{24}$$

for all $(\Delta\beta_1, \Delta\beta_2, \ldots, \Delta\beta_m)$ satisfying the constraint

$$F[(\beta_1, \beta_2, \ldots, \beta_m) + (\Delta\beta_1, \ldots, \Delta\beta_m)] = 0, \tag{25}$$

$F(0, \ldots, 0) = 0$ implies that $S(\beta_1, \beta_2, \ldots, \beta_m) \leqslant S(0, \ldots, 0) = 1$. Let us consider the vector

$$(\beta_1, \beta_2, \ldots, \beta_m) + \delta \quad \text{where } \delta = -\log[S(\beta_1, \beta_2, \ldots, \beta_m)] > 0. \tag{26}$$

We want to show that $(\beta_1, \beta_2, \ldots, \beta_m) + \delta$ is the solution of problem (17). That is, it satisfies

$$S[(\beta_1, \beta_2, \ldots, \beta_m) + \delta] = 1,$$

$$F[(\beta_1, \beta_2, \ldots, \beta_m) + \delta] \geqslant F[(\beta_1, \beta_2, \ldots, \beta_m) + \delta + (\Delta b_1, \ldots, \Delta b_m)] \tag{27}$$

for all $(\Delta b_1, \Delta b_2, \ldots, \Delta b_m)$ satisfying the constraint

$$S[(\beta_1, \beta_2, \ldots, \beta_m) + \delta + (\Delta b_1, \ldots, \Delta b_m)] = 1. \tag{28}$$

We first notice that functions $S(b_1, b_2, \ldots, b_m)$ and $F(b_1, b_2, \ldots, b_m)$ satisfy

$$S[(b_1, \ldots, b_m) + d] = S(b_1, \ldots, b_m)\exp(d),$$

$$F[(b_1, \ldots, b_m) + d] = F(b_1, \ldots, b_m) + d. \tag{29}$$

Applying relation (29), we see that $(\beta_1, \beta_2, \ldots, \beta_m) + \delta$ satisfies

$$S[(\beta_1, \beta_2, \ldots, \beta_m) + \delta] = S(\beta_1, \beta_2, \ldots, \beta_m)\exp(\delta) = 1,$$

$$F[(\beta_1, \beta_2, \ldots, \beta_m) + \delta] = F(\beta_1, \beta_2, \ldots, \beta_m) + \delta = \delta.$$

We prove (27) by contradiction. Suppose there is a $(\Delta b_1, \Delta b_2, \ldots, \Delta b_m)$ satisfying (28) and

$$F[(\beta_1, \beta_2, \ldots, \beta_m) + \delta + (\Delta b_1, \ldots, \Delta b_m)] > F[(\beta_1, \beta_2, \ldots, \beta_m) + \delta].$$

Consider the vector $(\beta_1, \beta_2, \ldots, \beta_m) + \delta + (\Delta b_1, \Delta b_2, \ldots, \Delta b_m) + d$ where

$$d = -F[(\beta_1, \beta_2, \ldots, \beta_m) + \delta + (\Delta b_1, \ldots, \Delta b_m)] < -F[(\beta_1, \beta_2, \ldots, \beta_m) + \delta] = -\delta. \tag{30}$$

Using relation (29), constraint (28), the definition of $\delta$ in (26), and inequality (30) we obtain

$$F[(\beta_1, \beta_2, \ldots, \beta_m) + \delta + (\Delta b_1, \ldots, \Delta b_m) + d] = F[(\beta_1, \beta_2, \ldots, \beta_m) - \delta + (\Delta b_1, \ldots, \Delta b_m)] + d = 0,$$

$$S[(\beta_1, \beta_2, \ldots, \beta_m) + \delta + (\Delta b_1, \ldots, \Delta b_m) + d] = S[(\beta_1, \beta_2, \ldots, \beta_m) + \delta + (\Delta b_1, \ldots, \Delta b_m)]\exp(d) < \exp(-\delta) = S(\beta_1, \beta_2, \ldots, \beta_m),$$

which contradicts with the definition of $(\beta_1, \beta_2, \ldots, \beta_m)$ that $S(b_1, b_2, \ldots, b_m)$ attains the constrained minimum at $(\beta_1, \beta_2, \ldots, \beta_m)$. Therefore, (27) must be true.

## References

Abbondanzieri, E.A., Greenleaf, W.J., Shaevitz, J.W., Landick, R., Block, S.M., 2005. Direct observation of base-pair stepping by RNA polymerase. Nature 438, 460–465.

Abrahams, J., Leslie, A., Lutter, R., Walker, J., 1994. Structure at 2.8 Å resolution of $F_1$-ATPase from bovine heart mitochondria. Nature 370, 621–628.

Astumian, R., 1997. Thermodynamics and kinetics of a Brownian motor. Science 276, 917–922.

Astumian, R.D., Derenyi, I., 1998. Fluctuation driven transport and models of molecular motors and pumps. Eur. J. Biophys. 27, 474–489.

Berg, H.C., 1993. Random Walks in Biology. Princeton University Press, Princeton, N.J.

Berg, H.C., 2003. The rotary motor of bacterial flagella. Ann. Rev. Biochem. 72, 19–54.

Block, S.M., Berg, H.C., 1984. Successive incorporation of force-generating units in the bacterial rotary motor. Nature 309, 470–473.

Block, S.M., Asbury, C.L., Shaevitz, J.W., Lang, M.J., 2003. Probing the kinesin reaction cycle with a 2D optical force clamp. Proc. Natl Acad. Sci. USA 100, 2351–2356.

Boyer, P., 1993. The binding change mechanism for ATP synthase–some probabilities and possibilities. Biochim. Biophys. Acta 1140, 215–250.

Boyer, P., 1997. The ATP synthase—a splendid molecular machine. Ann. Rev. Biochem. 66, 717–749.

Brockwell, P.J., Davis, R.A., 1998. Time Series: Theory and Methods. Springer, New York.

Coppin, C., Pierce, D., Hsu, L., Vale, R., 1997. The load dependence of kinesin's mechanical cycle. Proc. Natl. Acad. Sci. USA 94, 8539–8544.

Dimroth, P., Wang, H., Grabe, M., Oster, G., 1999. Energy transduction in the sodium *F*-ATPase of propionigenium modestum. Proc. Natl Acad. Sci. USA 96, 4924–4929.

Doering, C., Horsthemke, W., Riordan, J., 1994. Nonequilibrium fluctuation-induced transport. Phys. Rev. Lett. 72, 2984–2987.

Elston, T., Wang, H., Oster, G., 1998. Energy transduction in ATP synthase. Nature 391, 510–514.

Fan, J., Yao, Q., 2005. Nonlinear Time Series: Nonparametric and Parametric Methods. Springer, New York.

Fisher, R.A., 1922. On the mathematical foundations of theoretical statistics. Philos. Trans. R. Soc. London Ser. A 222, 309–368.

Fletcher, R., Reeves, C.M., 1964. Function minimization by conjugate gradients. Comput. J. 7, 148–154.

Gardiner, C., 1985. Handbook of Stochastic Methods, second ed. Springer, New York.

Harris, J.W., Stocker, H., 1988. Handbook of Mathematics and Computational Science. Springer, New York.

Hirono-Hara, Y., Ishizuka, K., Kinosita, K., Yoshida, M., Noji, H., 2005. Activation of pausing $F_1$ motor by external force. Proc. Natl Acad. Sci. USA 120, 4288–4293.

Hoel, P.G., 1962. Introduction to Mathematical Statistics. Wiley, New York.

Howard, J., Hudspeth, A.J., Vale, R.D., 1989. Movement of microtubules by single kinesin molecules. Nature 342, 154–158.

Hunt, A.J., Gittes, F., Howard, J., 1994. The force exerted by a single kinesin molecule against a viscous load. Biophys. J. 67, 766–781.

Itoh, H., Takahashi, A., Adachi, K., Noji, H., Yasuda, R., Yoshida, M., Kinosita, K., 2004. Mechanically driven ATP synthesis by $F_1$-ATPase Nature 427, 465–468.

Julicher, F., Ajdari, A., Prost, J., 1997. Modeling molecular motors. Rev. Mod. Phys. 69, 1269–1281.

Landau, L.D., Lifshitz, E.M., Pitaevskii, L.P., 1980. Statistical Physics, third rev. ed. Pergamon Press, Oxford, New York.

Menz, R.I., Walker, J.E., Leslie, A.G., 2001. Structure of bovine mitochondrial $F_1$-ATPase with nucleotide bound to all three catalytic sites: implications for the mechanism of rotary catalysis. Cell 106, 331–341.

Mogilner, A., Oster, G., 1999. The polymerization ratchet model explains the force–velocity relation for growing microtubules. Eur. J. Biophys. 28, 235–242.

Noji, H., Yasuda, R., Yoshida, M., Kinosita, K., 1997. Direct observation of the rotation of $F_1$-ATPase. Nature 386, 299–302.

Oster, G., Wang, H., 2000. Reverse engineering a protein: the mechanochemistry of ATP synthase. Biochim. Biophys. Acta (Bioenerg.) 1458, 482–510.

Oster, G., Wang, H., 2003. Rotary protein motors. Trends Cell Biol. 13, 114–121.

Peskin, C.S., Odell, G.M., Oster, G., 1993. Cellular motions and thermal fluctuations: the Brownian ratchet. Biophys. J. 65, 316–324.

Peskin, C.S., Ermentrout, G.B., Oster, G.F., 1994. Mechanochemical Coupling in ATPase Motors. Springer, Les Houches.

Polak, E., Ribiere, G., 1969. Note sur la convergence de methode de directions conjuguees. Rev. Fr. Inf. Rech. Oper. 16, 35–43.

Prost, J., Chauwin, J., Peliti, L., Ajdari, A., 1994. Asymmetric pumping of particles. Phys. Rev. Lett. 72, 2652–2655.

Reif, F., 1965. Fundamentals of Statistical and Thermal Physics. McGraw-Hill, New York.

Reimann, P., 2002. Brownian motors: noisy transport far from equilibrium. Phys. Rep. 361, 57–265.

Risken, H., 1989. The Fokker–Planck Equation, second ed. Springer, New York.

Sabbert, D., Engelbrecht, S., Junge, W., 1996. Intersubunit rotation in active *F*-ATPase. Nature 381, 623–625.

Sambongi, Y., Iko, Y., Tanabe, M., Omote, H., Iwamoto-Kihara, A., Ueda, I., Yanagida, T., Wada, Y., Futai, M., 1999. Mechanical rotation of the *c* subunit oligomer in ATP synthase ($F_0F_1$): direct observation. Science 286, 1722–1724.

Samuel, A., Berg, H., 1995. Fluctuation analysis of rotational speed of the bacterial flagellar motor. Proc. Natl Acad. Sci. 92, 3502–3506.

Schnitzer, M.J., Block, S.M., 1997. Kinesin hydrolyses one ATP per 8-nm step. Nature 388, 386–390.

Vale, R., 1986. Kinesin: possible biological roles for a new microtubule motor. TIBS 11, 464–468.

Visscher, K., Schnitzer, M., Block, S., 1999. Single kinesin molecules studied with a molecular force clamp. Nature 400, 184–189.

Walton, D.B., 2002. Analysis of single-molecule kinesin assay data by hidden Markov model filtering. Ph.D. Dissertation, University of Arizona.

Wang, H., 2003. Mathematical theory of molecular motors and a new approach for uncovering motormechanism. IEE Proc. Nanobiotechnol. 150, 127–133.

Wang, H., Oster, G., 1998. Energy transduction in the $F_1$ motor of ATP synthase. Nature 396, 279–282.

Wang, H., Oster, G., 2002. The Stokes efficiency for molecular motors and its applications. Europhys. Lett. 57, 134–140.

Wang, H., Peskin, C., Elston, T., 2003. A robust numerical algorithm for studying biomolecular transport processes. J. Theor. Biol. 221, 491–511.

Weber, J., Senior, A.E., 1997. Catalytic mechanism of $F_1$-ATPase. Biochim. Biophys. Acta 1319, 19–58.

Yasuda, R., Noji, H., Kinosita, K., Yoshida, M., 1998. $F_1$-ATPase is a highly efficient molecular motor that rotates with discrete 120° steps. Cell 93, 1117–1124.

Yasuda, R., Noji, H., Yoshida, M., Kinosita, K., Itoh, H., 2001. Resolution of distinct rotational substeps by submillisecond kinetic analysis of $F_1$-ATPase. Nature 410, 898–904.