Exact dynamic properties of molecular motors

N. J. Boon and R. B. Hoyle

Citation: The Journal of Chemical Physics **137**, 084102 (2012); doi: 10.1063/1.4746392
View online: http://dx.doi.org/10.1063/1.4746392
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/137/8?ver=pdfcov
Published by the AIP Publishing

Exact dynamic properties of molecular motors

N. J. Boon[a] and R. B. Hoyle[b]
Department of Mathematics, University of Surrey, Guildford GU2 7XH, United Kingdom

(Received 16 May 2012; accepted 2 August 2012; published online 23 August 2012)

Molecular motors play important roles within a biological cell, performing functions such as intracellular transport and gene transcription. Recent experimental work suggests that there are many plausible biochemical mechanisms that molecules such as myosin-V could use to achieve motion. To account for the abundance of possible discrete-stochastic frameworks that can arise when modeling molecular motor walks, a generalized and straightforward graphical method for calculating their dynamic properties is presented. It allows the calculation of the velocity, dispersion, and randomness ratio for any proposed system through analysis of its structure. This article extends work of King and Altman ["A schematic method of deriving the rate laws of enzyme-catalyzed reactions," J. Phys. Chem. **60**, 1375–1378 (1956)] on networks of enzymatic reactions by calculating additional dynamic properties for spatially hopping systems. Results for n-state systems are presented: single chain, parallel pathway, divided pathway, and divided pathway with a chain. A novel technique for combining multiple system architectures coupled at a reference state is also demonstrated. Four-state examples illustrate the effectiveness and simplicity of these methods. © *2012 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4746392]

I. INTRODUCTION

Directional intracellular transport, gene transcription and cell division are examples of important molecular processes that all living organisms require in order to function. Many different biological motors fulfill these roles at a molecular level through the transformation of chemical energy from ATP hydrolysis into mechanical work. This process occurs through a set of sequences of biochemical reactions and coordinated mechanical motions. Biomechanochemical pathways of varying complexity have been suggested for many molecules, for example, for myosin-V.[1–8]

Single molecule measurements can determine average dynamic quantities of molecular motors such as their velocity as well as their fluctuations or dispersion about these averages. Such results can reveal much about the underlying biomechanochemical pathways which are difficult to measure directly. Analytical techniques for calculating these quantities are important tools to bridge the gap between measured variables and postulated processes. This paper presents novel methods for calculating the velocity, dispersion, and randomness ratio of processive molecular motors from the underlying quantities in these processes. The focus here is on molecular motors but these methods apply to anything that hops along a periodic lattice in physical space.

A discrete stochastic model of a molecular motor walk assumes that the biomechanochemical pathways can be split into discrete states and that transitions between these states occur probabilistically. It is assumed to approach its steady state (equilibrium in time) rapidly if molecular detachments are neglected. The steady-state solution of the governing master equations therefore can be used to determine the relevant mean behavior of the molecules, for example their average velocity or the dispersion, from the transition rates. This is important in the investigation of many molecular motor models, for example, those for myosin-V.[1,3,7]

The flux balance method[9] allows calculation of quantities such as the velocity or the dwell times, without the need for explicit solutions for the state probabilities. However, it cannot give quantities such as the dispersion or randomness ratio, the reciprocal of which is the number of rate-limiting steps.[10] A method presented by Chemla *et al.*[11] allows the calculation of velocities, dispersions for any given biochemical pathway but cannot give general formulae. The calculations, particularly for large systems with reversible transitions, can require computationally expensive calculations and are mathematically quite involved.

An approach based on Derrida[12] has proved useful in calculating exact steady states and dynamic properties for specific classes of system architectures of arbitrary size. The simpler examples of these include single chains,[12] parallel chains,[13] and divided pathways.[14] Periodic parallel lattices have also been studied[15] in the limit of strong coupling between each branch. Each class can be modified to include branches and molecular detachment.[16] The average velocity and its dispersion is calculated individually for each system architecture. The method presented here simplifies, consolidates, and extends all this work by presenting a general graphical method for any system.

A method for finding the steady-state probabilities of enzymatic networks graphically was first presented by King and Altman[17] and developed by Hill.[18] We have tailored it here to the context of molecular motors and extended the method to give additional dynamic quantities such as the dispersion presented in Sec. III, that was previously difficult to calculate.

[a]n.boon@surrey.ac.uk.
[b]r.hoyle@surrey.ac.uk.

A novel and mathematically straightforward method for calculating dynamic quantities for any biochemical pathway with any distribution of stepping sizes is presented. Explicit and exact expressions for the steady-state probabilities, average velocity, and dispersion relation are given. These all depend on a set of variables C_{ij} that can be determined in an intuitive graphical manner from the system architecture. The ability to calculate results for any generalized structure in such a straightforward manner distinguishes this method from all others; these methods enable the derivation of general formulae for specific system structures reducing potentially expensive calculations. The structure of the system is preserved in the calculations allowing existing general results to be analyzed and modified; we demonstrate a method of combining several general architectures together by coupling them at a reference state. For smaller systems the dispersion relation that is usually complicated can be written down simply thus reducing the level of mathematical complexity. The methods and expressions presented here are therefore powerful tools in investigating the steady states and dynamic properties of theoretical models for molecular motor stepping cycles.

The calculation of steady-state probabilities and dynamic properties using this graphical method presented in Secs. II and III, respectively. An expression for the dispersion relation is given in terms of variables C_{ij} in Sec. III. Methods to simplify the calculation of the C_{ij} for large systems are shown in Sec. IV as well as a technique to obtain them from component structures coupled at a reference state. Examples of arbitrary-sized single chain, parallel pathway, and divided pathway are derived and extended to the novel divided pathway with a chain model using the combing technique in Sec. V. Example 4-state models also demonstrate these methods in Sec. VI. The paper concluded with a discussion in Sec. VII.

II. GENERAL STEADY-STATE PROBABILITIES

We consider a system of n states, each representing a biomechanochemical state of a molecule, described by n master equations whose form is determined by the proposed set of biochemical pathways. In matrix form, we have

$$\dot{\mathbf{P}}(t) = \mathbb{M}\mathbf{P}(t), \tag{1}$$

where $\mathbb{M}$ is a $n \times n$ transition rate matrix and the ith component of the vector of state occupancy probability $\mathbf{P}(t)$ is $P_i(t)$. This can be written as

$$\frac{dP_i}{dt} = \sum_j [W_{ji}P_j - (W_{ij} + \delta_i)P_i], \tag{2}$$

where the transition rate from state i to state j is denoted by W_{ij} and the detachment rate (the rate at which molecules leave the pathway) from state i is given by δ_i. We have $W_{ij} = 0$ if there is no possible transition between state i and state j. A system with non-zero detachment rates can be *renormalized* into a system without detachments using a procedure outlined by Kolomeisky and Fisher.[16] Therefore, only systems without molecular detachment ($\delta_i = 0$) are considered in this article.

Thus in the steady state,

$$\sum_j [W_{ji}P_j - P_i W_{ij}] = 0 \forall j. \tag{3}$$

A molecular motor can be assumed to pass through a repeating sequence of biomechanochemical changes to achieve motion. There is therefore a periodicity to the system with the transition from one period to another carrying some notion of direction: *forwards* moves to the next period and *backwards* moves to the previous. For example, the master equations for states along a single chain (nearest-neighbor coupled states) are of the form

$$\frac{dP_i}{dt} = u_{i-1}P_{i-1} + w_{i+1}P_{i+1} - [u_i + w_i]P_i \tag{4}$$

with forwards transition rates denoted by u_i and backwards transition rates denoted by w_i and state $n \equiv 0$ and indices being taken modulo n. This is exactly the system studied by Derrida.[12]

We now introduce some useful definitions.

A *branch* is a sequence of states with only nearest-neighbor transitions between them.

A *coupling state* is a state that connects two or more branches.

A *rate path* from a to b is a product of rates along the directed path from a to b. For the system described in Eq. (4), $u_a u_{a+1} u_{a+2}$ is a rate path from a to $a + 3$. A rate path from a to b is *closed* if it also contains a rate path from b to a.

A *rate tree* of b is a product of state-unique reaction rates (only one from each state) that contains a rate path from each state in itself to b.

A *configuration* of b is a non-unique rate tree of b containing one rate from every state in the system except b and a *configuration** of b is a non-unique rate tree of b containing one rate from every state including b. A configuration cannot contain any closed rate paths; however, a configuration* of b must contain exactly one closed rate path.

In the single chain $n = 4$ example above, a configuration of state 1 is $u_2u_3u_0$, another configuration is $w_2u_3u_0$. The remaining two are $w_2w_3u_0$ and $w_2w_3w_0$. This is shown in Figure 1. Note that $u_2u_3w_0$ is not a configuration of state 1 because it does not contain a rate path from 0 or 3 to 1. In this example, a configuration* of state 1 is simply a configuration multiplied by either u_1 or w_1.

A *rate path reversal* is a rate path within a rate tree with forwards rates changed into backwards rates so that the rate tree retains its properties. In the previous example, $w_2u_3u_0$

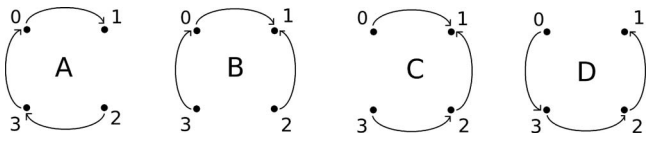

FIG. 1. All the configurations of state 1 for the Derrida[12] $n = 4$ example. A represents configuration $c_A = u_2u_3u_0$, B represents $c_B = w_2u_3u_0$, C represents $c_C = w_2w_3u_0$, and $c_D = w_2w_3w_0$. The steady-state solution for state 1 is therefore $P_1/N = Q_1 = c_A + c_B + c_C + c_D$. The configurations* of state 1 are given by u_1c_i or w_1c_i with $i \in \{A, B, C, D\}$.

and $w_2w_3u_0$ are rate path reversals of $u_2u_3u_0$ but $u_2u_3w_0$ is not.

If Q_i is the sum of all possible configurations for state i and Q_i^* is the sum of all possible configurations* for state i, then

$$Q_i^* = \sum_j W_{ji} Q_j \tag{5}$$

and

$$Q_i^* = Q_i \sum_j W_{ij}, \tag{6}$$

with both giving a relation between a sum over all configurations* of i and a sum over all configurations of i. The first relation being a sum over each state j of Q_j multiplied by the reaction rate from j to i, gives a rate tree for i that contains a rate from every state—a sum over all configurations* of i. The second being a sum over all configurations of i state multiplied by a sum over rates from i to every other state, also gives a sum over all configurations* of i.

Thus

$$\sum_j [W_{ji} Q_j - W_{ij} Q_i] = 0 \,\forall i. \tag{7}$$

In the single chain $n = 4$ example, taking indices modulo 4,

$$\begin{aligned}
\sum_j W_{ji} Q_j &= u_{i-1} Q_{i-1} + w_{i+1} Q_{i+1}, \\
&= u_{i-1}(u_i u_{i+1} u_{i+2} + w_i u_{i+1} u_{i+2}) \\
&\quad + u_{i-1}(w_i w_{i+1} u_{i+2} + w_i w_{i+1} w_{i+2}) \\
&\quad + w_{i+1}(u_{i-1} u_i u_{i+2} + u_{i-1} u_i w_{i+2}) \\
&\quad + w_{i+1}(w_{i-1} u_i w_{i+2} + w_{i-1} w_i w_{i+2}), \\
&= (u_i + w_i)(u_{i-1} u_{i+1} u_{i+2} + u_{i-1} w_{i+1} u_{i+2} \\
&\quad + u_{i-1} w_{i+1} w_{i+2} + w_{i-1} w_{i+1} w_{i+2}), \\
&= Q_i(u_i + w_i), \\
&= Q_i \sum_j W_{ij}.
\end{aligned}$$

Equation (7) shows that the Q_i satisfy Eq. (3). Therefore, the sum of all possible configurations for state i is the non-normalized steady-state probability for state i and so

$$P_i = N Q_i, \tag{8}$$

where N is a normalization constant that ensures the probabilities sum to unity. This is the result first shown by King and Altman[17] and developed by Hill.[18]

In the steady-state, the equations investigated by Derrida[12] become

$$0 = u_{i-1} P_{i-1} + w_{i+1} P_{i+1} - [u_i + w_i] P_i, \tag{9}$$

and therefore have the solution

$$\begin{aligned}
\frac{P_i}{N} &= \prod_{j=i+1}^{i-1} u_j + w_{i+1} \prod_{j=i+2}^{i-1} u_j + \cdots + \prod_{j=i+1}^{i-1} w_j, \\
\frac{P_i}{N e_1} &= \frac{1}{u_i} \left(1 + \sum_{j=i+1}^{i-1} \prod_{i=i+1}^{j} \frac{w_i}{u_i} \right),
\end{aligned} \tag{10}$$

a forwards rate path from $i+1$ to i (modulo n) plus all its path reversals, where $e_1 = \prod_{j=1}^{n-1} u_j$. This is exactly the solution shown by Derrida.[12]

The probabilistic steady-state can be found using this method for any closed system. Physically, this informs us as to where in the biochemical pathways the molecules tend to dwell.

III. DYNAMIC PROPERTIES

The probabilistic steady state allows the calculation of the dynamic properties of the system. Again we have a periodic system with n states and a rate from state i to state j is denoted by W_{ij}. However, now physical distances between states must also be specified. State i is a distance d_i from reference state 0 and the total physical distance over the whole period is d (Figure 2).

We want to calculate the average velocity v and the dispersion D of molecules in the system and so we consider the movement of molecules in physical space along a periodically repeating lattice of physical sites. The probability of being in the ith site on the sth cycle is denoted by $p_{i,s}$. Each site is connected to $n-1$ sites forwards and $n-1$ sites backwards, assuming that a molecule cannot jump a cycle length or longer for simplicity. For example, site $(0, s)$ is connected to each site (i, s) and $(i, s-1)$ for all $i \neq 0$. The forwards and backwards transition rates from site i to site j are u_{ij} and w_{ij}, respectively. The distance from site $(0, s)$ to site (i, s) is denoted by d_i and $d_0 = 0$. This system is shown in Figure 2.

The site occupancy probabilities $p_{i,s}$ are given by

$$\begin{aligned}
\frac{dp_{i,s}}{dt} &= \sum_{j=i+1}^{n-1} (w_{ji} p_{j,s} + u_{ji} p_{j,s-1}) \\
&\quad + \sum_{j=0}^{i-1} (w_{ji} p_{j,s+1} + u_{ji} p_{j,s}) \\
&\quad - \sum_j (u_{ij} + w_{ij}) p_{i,s}.
\end{aligned} \tag{11}$$

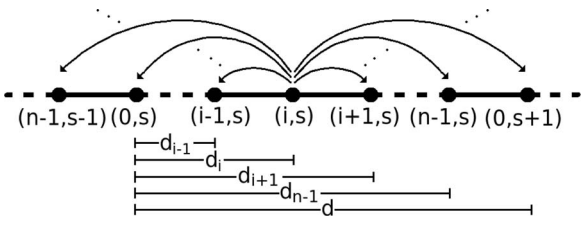

FIG. 2. The periodic 1D lattice of sites in physical space, ordered in terms of their distances (labeled d_i) from the reference state $(0, s)$. The possible site transitions from state (i, s) are shown by the arrows.

Here each s is associated with one repeat of the physical-space lattice, and $\sum_s p_{i,s} = P_i$. Note that when $i = 0$, $\sum_{j=0}^{i-1}$ is defined to give 0 and similarly when $i = n-1$ for $\sum_{j=i+1}^{n-1}$.

A. Velocity

The average displacement of a molecule along the track is given by

$$\begin{aligned}\langle x\rangle &= \sum_s \sum_i p_{i,s}(ds + d_i)\\ &= \sum_i (dX_i + d_i P_i),\end{aligned} \tag{12}$$

where $X_i \equiv \sum_s s p_{i,s}$.

Assuming that the system is in its steady state, we have $\frac{dP_i}{dt} = 0$ and so

$$\begin{aligned}v &= \frac{d\langle x\rangle}{dt},\\ &= d\sum_i \frac{dX_i}{dt}.\end{aligned} \tag{13}$$

Therefore, Eq. (A3) gives that

$$v = d\sum_i \sum_{j=0}^{i-1} (u_{ji}P_j - w_{ji}P_j), \tag{14}$$

exactly as expected from flux balance.[9] Note the velocity is independent of the d_i and only depends on the total step size d.

B. Dispersion

The dispersion is defined to be

$$D \equiv \frac{1}{2}\lim_{t\to\infty}\frac{d}{dt}(\langle x^2\rangle - \langle x\rangle^2). \tag{15}$$

Therefore, from Eq. (A23)

$$\begin{aligned}D = {} & d^2 \sum_i \sum_{j=0}^{i-1} \frac{1}{2}(u_{ji} + w_{ji})P_j\\ & + \frac{d^2}{Q_0}\sum_i \sum_{j=0}^{i-1}[w_{ji} - u_{ji}]\sum_{k\neq 0} G_k C_{jk}\\ & + \frac{dv}{Q_0}\sum_i \sum_{j\neq 0} G_j C_{ij},\end{aligned} \tag{16}$$

with

$$G_i = \frac{v}{d}P_i - \sum_{j=i+1}^{n-1} u_{ji}P_j + \sum_{j=0}^{i-1} w_{ji}P_j, \tag{17}$$

and the C_{ij} are the sum over all configurations of state $i \neq 0$ given a non-zero rate from 0 to j and divided by that rate. It can be deduced from this definition that

$$Q_0 = \sum_{j\neq 0} W_{j0}C_{jk},\ \forall k \neq 0, \tag{18}$$

$$Q_i = \sum_{k\neq 0} W_{0k}C_{ik}\ i \neq 0, \tag{19}$$

which give the steady-state probabilities P_i once normalized.

Equations (14) and (16) are general under the assumptions that the states lie on a 1D physical lattice and that it is not possible for a molecule cannot jump a cycle length or longer. Only the C_{ij} need to be calculated for each individual system as general explicit equations have been derived in terms of them for the state-occupancy probabilities (Eqs. (18) and (19)), the velocity (Eq. (14)), and the dispersion (Eq. (16)). Systems with high degrees of symmetry can greatly simplify these calculations as shown in Sec. IV.

The randomness ratio is given by

$$r = \frac{2D}{dv}. \tag{20}$$

Note that assuming that the transition rates are independent of the substeps d_i, the velocity, dispersion, and therefore randomness ratio are also.

IV. CALCULATING THE C_{ij}

Equations (14) and (16) give general expressions for the velocity and the dispersion respectively for any 1D hopping system assuming a molecule cannot jump one repeat or more of the lattice of physical sites. The C_{ij} must be derived for any individual system structure; however, analysis of the architecture can greatly simplify the calculation.

Each term of a given C_{ij} must obey three rules. First, it must contain a rate path from j to i not through 0. Second, for any state $a \neq 0, i, j$ it must contain exactly one rate path from a to i or a to 0. Third, it must contain exactly one transition rate from each state except 0 and i—from which there should be none.

For illustrative purposes, the C_{ij} for the three smallest completely general systems are given. The $n = 2$ system has

$$C_{11} = 1,$$

the $n = 3$ system has

$$\begin{aligned}C_{11} &= W_{20} + W_{21},\\ C_{12} &= W_{21},\\ C_{21} &= W_{12},\\ C_{22} &= W_{10} + W_{12},\end{aligned}$$

and the $n = 4$ system has

$$\begin{aligned}C_{11} = {} & (W_{20} + W_{21})(W_{30} + W_{31} + W_{32})\\ & + W_{23}(W_{30} + W_{31}),\end{aligned} \tag{21}$$

$$C_{12} = W_{21}(W_{30} + W_{31} + W_{32}) + W_{23}W_{31}, \tag{22}$$

$$C_{13} = W_{31}(W_{20} + W_{21} + W_{23}) + W_{32}W_{21}, \tag{23}$$

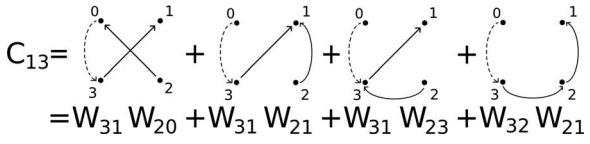

FIG. 3. A graphical example of how C_{13} is calculated for the $n = 4$ system. Each rate W_{ij} included in C_{13} is shown as a solid arrow. C_{13} is the sum over all configurations of state 1 given a non-zero rate from 0 to 3 shown as a dashed arrow. The algebraic interpretation is displayed below the graphical one. Note that to satisfy the definition of a configuration of 1 there is a rate path from any state to state 1. Each C_{ij} can be interpreted in this manner.

$$C_{21} = W_{12}(W_{30} + W_{31} + W_{32}) + W_{13}W_{32}, \tag{24}$$

$$C_{22} = (W_{10} + W_{12})(W_{30} + W_{31} + W_{32}) + W_{13}(W_{32} + W_{30}), \tag{25}$$

$$C_{23} = W_{32}(W_{12} + W_{10} + W_{13}) + W_{31}W_{12}, \tag{26}$$

$$C_{31} = W_{13}(W_{20} + W_{21} + W_{23}) + W_{12}W_{23}, \tag{27}$$

$$C_{32} = W_{23}(W_{10} + W_{12} + W_{13}) + W_{21}W_{13}, \tag{28}$$

$$C_{33} = (W_{10} + W_{13})(W_{20} + W_{21} + W_{23}) + W_{12}(W_{20} + W_{23}). \tag{29}$$

The calculation of C_{13} is explained as an example in Figure 3.

It is also possible to consider arbitrary-sized systems in terms of branched states and coupling states. Considering a branch k of states with nearest neighbor interactions and defining u_i^k as the rate from i to $i + 1$ and w_i^k as the rate from i to $i - 1$, the following notation is useful when writing down configurations:

$$ {}^k\Lambda_a^b \equiv \prod_{i=a}^{b} u_i^k, \tag{30}$$

$$ {}^k\Pi_a^b \equiv \prod_{i=a}^{b} w_i^k, \tag{31}$$

$$ {}^k\Xi_a^b \equiv {}_k\Lambda_a^b \left(1 + \sum_{j=a}^{b} \prod_{i=a}^{j} \frac{w_i^k}{u_i^k}\right), \tag{32}$$

and define that each expression becomes unity if no rates are included, for example ${}^k\Xi_a^{a-1} = 1$.

${}^k\Lambda_a^b$ and ${}^k\Pi_a^b$ represent a path and a reversed rate path, respectively, between a and b on branch k. ${}^k\Xi_a^b$ represents the sum of all possible rate path reversals between states a and b on branch k. This is shown graphically in Figure 4. Multiplying several of these components together and ensuring the indices do not overlap gives the properties of all the terms. For example, ${}^k\Xi_a^b\,{}^k\Xi_{b+1}^c$ represents the sum of all combinations of path reversals between a and b and between $b + 1$ and c on branch k. ${}^k\Xi_a^b\,{}^{k'}\Xi_c^d$ represents the sum of all combinations of path reversals between a and b on branch k and all combinations of path reversals between c and d on branch k'.

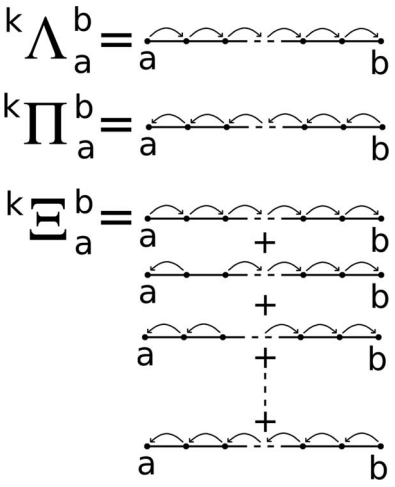

FIG. 4. A graphical representation of ${}^k\Lambda_a^b$ the rate path from a to b, ${}^k\Pi_a^b$ the reversed rate path from b to a, and ${}^k\Xi_a^b$ the sum of all rate path reversals between a and b all along a branch k.

A given C_{ij} is written in terms of the rates from coupling states and rates from states on a branch. Coupling-state rates appear explicitly in the equations, while branch-state rates can be grouped together using relations (30), (31) and (32). For example, considering the simplest system architecture of only one branch and no coupling states known as the single chain (Figure 5(a)), we have

$$C_{ij} = u_j \Lambda_{j+1}^{i-1} \Xi_1^{j-1} \Xi_{i+1}^{n-1}, \; i > j, \tag{33}$$

$$C_{ij} = w_j \Pi_{i+1}^{j-1} \Xi_1^{i-1} \Xi_{j+1}^{n-1}, \; i < j, \tag{34}$$

with the $i = j$ case given by Eq. (33) with $u_j \Lambda_{j+1}^{i-1} \to 1$. In this notation C_{ij} must be written as three separate equations because the relative locations of states i and j are important. A system with two branches and one coupling state defined to be the reference state 0, the parallel pathway (Figure 5(b)), has

$$C_{ij}^k = u_j^k\, {}^k\Lambda_{j+1}^{i-1}\, {}^k\Xi_1^{j-1}\, {}^k\Xi_{i+1}^{n^k-1}\, {}^{k'}\Xi_1^{n^{k'}-1}, \; i > j, \tag{35}$$

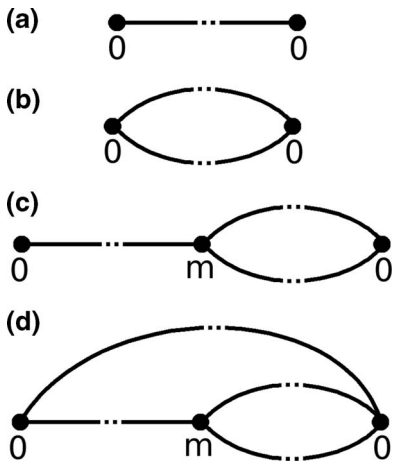

FIG. 5. A structural representation of (a) the single chain, (b) the two branch parallel pathway, (c) the divided pathway, and (d) the divided pathway with a chain coupled at state 0. Dots represent coupling states and lines represent branches.

$$C_{ij}^{k} = w_j^{k\,k}\Pi_{i+1}^{j-1\;k}\Xi_1^{i-1\;k}\Xi_{j+1}^{n^k-1\;k'}\Xi_1^{n^{k'}-1},\ i<j, \quad (36)$$

with $k \in \{1, 2\}$ and the $i = j$ case given by Eq. (35) with $u_j^{k\,k}\Lambda_{j+1}^{i-1} \to 1$. The case where the j, k state lies on a different branch to the i, k' state (i.e., $k \neq k'$) has $C_{(i,k)(j,k')} = 0$ because no configurations of i given a rate $W_{0,(j,k)}$ exist as there is no rate path from j, k to i, k' that does not pass through 0.

Note that each expression can be derived from the other; in changing the order of i and j, $i \geq j \leftrightarrow i < j$, three transformations must be applied:

$$u_j^k \leftrightarrow w_j^k, \quad (37)$$

$${}^k\Xi_j^i,\ {}^k\Lambda_j^i,\ {}^k\Pi_j^i, \leftrightarrow {}^k\Xi_i^j,\ {}^k\Lambda_i^j,\ {}^k\Pi_i^j, \quad (38)$$

$${}^k\Lambda \leftrightarrow {}^k\Pi, \quad (39)$$

namely the rate from the jth state changes direction, the is and js in the indices of the grouped branch terms are swapped and Λ and Π are swapped. In this manner only a few expressions for the C_{ij} need to be written down, the rest can be deduced from these.

Introducing another coupling state adds another level of complexity. For a given C_{ij}, states i and j can now be this coupling state and not just branch states and so more than three expressions are required to fully define all the C_{ij}. For a system with two branches and two coupling states, defining one of the coupling states as the reference state 0 and denoting the other by m we have the divided pathway,[14] Figure 5(c). For $0 < j < i < m$,

$$\begin{aligned} C_{ij} = u_j\Lambda_{j+1}^{i-1}\Xi_1^{j-1}\big[&w_m\Pi_{i+1}^{m-1\;k}\Xi_{m+1}^{n^k-1\;k'}\Xi_{m+1}^{n^{k'}-1} \\ &+ \Xi_{i+1}^{m-1}\big(u_m^{k\,k}\Lambda_{m+1}^{n^k-1\;k'}\Xi_{m+1}^{n^{k'}-1} \\ &+ u_m^{k'\,k'}\Lambda_{m+1}^{n^{k'}-1\;k}\Xi_{m+1}^{n^k-1}\big)\big], \end{aligned} \quad (40)$$

however it is much simpler to consider this in graphical form as in Figure 6. The case where $0 < j = i < m$ is recovered by sending $u_j^{k\,k}\Lambda_{j+1}^{i-1} \to 1$. The case $0 < i \leq j < m$ is given by applying transformations given by relations (37), (38) and (39) and an additional $u_m^k \leftrightarrow w_m$ transformation to take into account the extra coupling state.

In this manner all the C_{ij} for the general n-state divided pathway can be written down. This is presented in Appendix C.

A. Modifying system structures

The method discussed in this article preserves the structure of a given architecture within the calculations. This allows solutions for one architecture to be deduced from solutions to another, for example by modifying the structure of the system.

Modifying the single chain model into the parallel pathway model is akin to adding additional single chains into the system coupled at 0. Note that the parallel chain results given in Eqs. (35) and (36) are the single chain results multiplied by the product of each additional branch of the configuration of state 0 restricted to that branch.

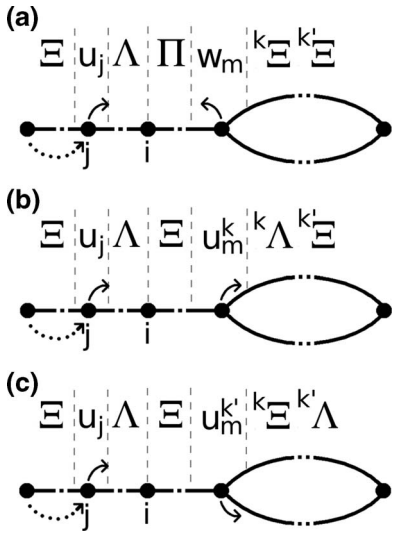

FIG. 6. A graphical representation of the C_{ij} for a divided pathway model for $0 < j < i < m$. Each figure part represents a term in Eq. (40) and contains all possible components of this C_{ij} where the coupling state rate points (a) backwards, (b) forwards along branch k, and (c) forwards along branch k'. Indices on the branched terms have been omitted—each symbol pertains to the branch states bracketed by the vertical dashed lines.

Therefore, defining ${}^SQ_0^k$ to be the sum of all configurations of 0 restricted to the single chain branch k and ${}^SC_{ij}^k$ to be the single chain C_{ij} restricted to branch k, gives the parallel pathway ${}^PC_{ij}^k$:

$${}^PC_{ij}^k = {}^SC_{ij}^k \prod_{k' \neq k} {}^SQ_0^{k'}. \quad (41)$$

Further results can be derived using these ideas. For example the divided pathway with an additional single chain coupled at 0 (Figure 5(d)) would have

$${}^{D+S}C_{ij} = {}^DC_{ij}\ {}^SQ_0,\ i, j \in \text{D section}, \quad (42)$$

$${}^{D+S}C_{ij} = {}^SC_{ij}\ {}^DQ_0,\ i, j \in \text{S section}, \quad (43)$$

$${}^{D+S}C_{ij} = 0,\ \text{otherwise}, \quad (44)$$

where D represents the divided pathway, S represents the single chain, ${}^DC_{ij}$, ${}^DQ_0$ are the C_{ij} and the Q_0, respectively, for just the divided pathway section and ${}^SC_{ij}$, ${}^SQ_0$ are the C_{ij} and the Q_0, respectively, for just the single chain section.

In general for two structures A and B coupled only at the reference state, the ${}^{A+B}C_{ij}$ for the combined structure can be written in terms of individual component structure variables

$${}^{A+B}C_{ij} = {}^AC_{ij}\ {}^BQ_0,\ i, j \in \text{A section}, \quad (45)$$

$${}^{A+B}C_{ij} = {}^BC_{ij}\ {}^AQ_0,\ i, j \in \text{B section}, \quad (46)$$

$${}^{A+B}C_{ij} = 0,\ \text{otherwise}. \quad (47)$$

In this manner one can derive results for the probabilistic steady state and dispersion for arbitrarily large and highly complex systems by splitting them into individual component systems coupled at the reference state and computing the C_{ij}

for each system individually. This is a very useful tool for investigating complex underlying biochemical pathways of a molecular motor.

V. GENERALIZED SYSTEM ARCHITECTURES

Many biologically inspired systems suggest that there is at least one state shared between all the forwards mechanochemical cycles, for example with myosin-V.[1–5,7,8] Defining this state to be 0 we can use this as a boundary of the periodic lattice of physical sites. Systems that do not have this property require a minimum of 3 branches and 4 coupling states and for simplicity we will not consider these here. Therefore, physical sites $(0, s)$ are connected to $2n-2$ other sites and sites $(i \neq 0, s)$ are connected to $n-1$ other sites. This is similar to the system in Figure 2 except sites $(i \neq 0, s)$ only have possible transitions to other sites $(j \neq i, s)$ and $(0, s+1)$. The equation for the velocity can now be simplified to

$$v = d \sum_{i \neq 0} \left[u_{i0} P_i - w_{0i} P_0 \right], \tag{48}$$

and the equation for the dispersion can also be simplified to

$$\begin{aligned} D = & \frac{d^2}{2} \sum_i (u_{i0} P_i + w_{0i} P_0) \\ & - \frac{d^2}{Q_0} \sum_i u_{i0} \sum_{j \neq 0} G_j C_{ij} \\ & + \frac{dv}{Q_0} \sum_i \sum_{j \neq 0} G_j C_{ij}, \end{aligned} \tag{49}$$

with

$$G_i = \frac{v}{d} P_i + w_{0i} P_0, \text{ for } i \neq 0. \tag{50}$$

From now on we will only consider systems of this type.

There are many different classes of n-state system architectures, each defined by the conditions on the transition rates.

A. Single chain model

A single chain model is a system with only nearest neighbor transitions as shown in Figure 7(a). The state occupancy probabilities P_i are governed by n master equations

$$\frac{dP_0}{dt} = u_{n-1} P_{n-1} + w_1 P_1 - (u_0 + w_0) P_0, \tag{51}$$

$$\frac{dP_1}{dt} = u_0 P_0 + w_2 P_2 - (u_1 + w_1) P_1, \tag{52}$$

$$\vdots$$

$$\begin{aligned} \frac{dP_{n-1}}{dt} = & u_{n-2} P_{n-2} + w_0 P_0 \\ & - (u_{n-1} + w_{n-1}) P_{n-1}. \end{aligned} \tag{53}$$

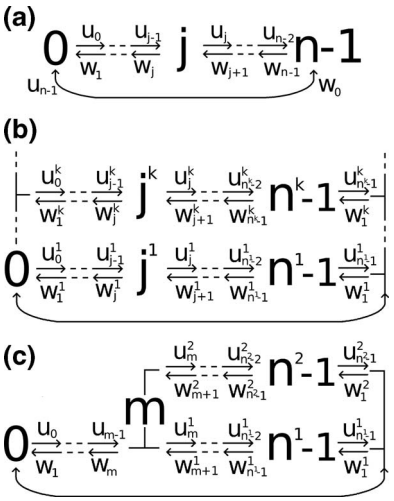

FIG. 7. (a) A n-state single chain system with n substeps and n different molecular detachment pathways. Forward rates are denoted by u_i and the backward rates by w_i. (b) A parallel pathway reaction network with an arbitrary number of branches. (c) A generalized divided pathway reaction network. Distances between states are not shown.

The steady state of the system is given by the methods in Sec. II. In this case it is exactly the result shown by Derrida[12]

$$Q_i = \frac{P_i}{N} = \frac{e_1}{u_i} \left(1 + \sum_{j=i+1}^{i-1} \prod_{i=i+1}^{j} \frac{w_i}{u_i} \right), \tag{54}$$

where N is the normalizing factor so that the probabilities sum to unity and

$$e_1 = \prod_{i=0}^{n-1} u_i. \tag{55}$$

Therefore,

$$Q_i = \Xi_{i+1}^{i-1}, \tag{56}$$

with indices being taken modulo n as state $0 \equiv n$.

Equation (48) gives the velocity of the system

$$v = d[u_{n-1} P_{n-1} - w_0 P_0], \tag{57}$$

and Eq. (49) gives the dispersion of the system

$$\begin{aligned} D = & \frac{d^2}{2} (u_{n-1} P_{n-1} + w_0 P_0) \\ & - \frac{d^2}{Q_0} u_{n-1} \sum_{j \neq 0} G_j C_{n-1,j} \\ & + \frac{dv}{Q_0} \sum_i \sum_{j \neq 0} G_j C_{ij}, \end{aligned} \tag{58}$$

with

$$G_i = \frac{v}{d} P_i + w_0 P_0, \text{ for } i \neq 0, \tag{59}$$

and

$$C_{ij} = \Lambda_j^{i-1}\Xi_1^{j-1}\Xi_{i+1}^{n-1},\ i \geq j, \tag{60}$$

$$C_{ij} = \Pi_{i+1}^{j}\Xi_1^{i-1}\Xi_{j+1}^{n-1},\ i < j. \tag{61}$$

This is exactly the result shown by Derrida.[12]

B. Parallel pathway model

The parallel pathway model is a simple modification to the single chain model. In this section we demonstrate how the single chain solution can be modified to give us the parallel pathway result. For the parallel pathway model with an arbitrary number of branches (enumerated by superscript k), the architecture is shown in Figure 7(b) and the steady state is given by

$$Q_0 = \frac{P_0}{N} = \sum_k {}^k\Xi_1^{n^k-1} \tag{62}$$

$$Q_i^k = \frac{P_i^k}{N} = {}^k\Xi_{i+1}^{i-1}\sum_{k'\neq k} {}^{k'}\Xi_1^{n^{k'}-1}. \tag{63}$$

Equation (48) gives the velocity

$$v = d\sum_k \left[u_{n^k-1}^k P_{n^k-1}^k - w_0^k P_0\right]. \tag{64}$$

The dispersion is given by Eq. (49)

$$\begin{aligned} D = {} & \frac{d^2}{2}\sum_k \left(u_{n^k-1}^k P_{n^k-1}^k + w_0^k P_0\right) \\ & - \frac{d^2}{Q_0}\sum_k u_{n^k-1}^k \sum_{j\neq 0} G_j^k C_{n^k-1,j}^k \\ & + \frac{dv}{Q_0}\sum_k\sum_i\sum_{j\neq 0} G_j^k C_{ij}^k, \end{aligned} \tag{65}$$

with

$$G_i^k = \frac{v}{d}P_i^k + w_0^k P_0^k,\ \text{for } i \neq 0, \tag{66}$$

where

$$C_{ij}^k = {}^k\Lambda_j^{i-1}\ {}^k\Xi_1^{j-1}\ {}^k\Xi_{i+1}^{n^k-1}\prod_{k'\neq k} {}^{k'}\Xi_1^{n^{k'}-1}, \tag{67}$$

for $i \geq j$, and

$$C_{ij}^k = {}^k\Pi_{i+1}^{j}\ {}^k\Xi_1^{i-1}\ {}^k\Xi_{j+1}^{n^k-1}\prod_{k'\neq k} {}^{k'}\Xi_1^{n^{k'}-1} \tag{68}$$

for $i < j$.

Restricting to only two branches, $k \in [1, 2]$, the resulting dispersion gives exactly the result shown by Kolomeisky.[13] Our result is general for any number of parallel branches.

C. Divided pathway and divided pathway with a chain

The generalized divided pathway is shown in Figure 7(c). The divided pathway with a chain is constructed by coupling a single chain to the reference state shown schematically in Figure 5(d).

The velocity, dispersion, and the G_i^k relations for both of these systems are the same as those for the parallel pathway system (although the C_{ij} are different) and are given by Eqs. (64), (65) and (66), respectively. For the divided pathway $k \in \{1, 2\}$ where each k represents a different divided branch and the C_{ij} are given in Appendix C. For the divided pathway with a chain $k \in \{1, 2, 3\}$ where the additional $k = 3$ branch represents the single chain and the C_{ij} are derived from those of the single chain and the divided pathway using the method given in Sec. IV A.

The steady-state probabilities are calculated from the C_{ij} using Eqs. (18) and (19).

VI. FOUR STATE MODEL EXAMPLE

Biologically interesting models for molecular motors exist that involve relatively few states.[1,3,7,19] Our more intuitive framework gives the dispersion much more readily than existing approaches for systems with smaller number of states. The generalized $n = 4$ system is shown in Figure 8 and has unnormalized probabilities

$$Q_0 = W_{10}C_{1k} + W_{20}C_{2k} + W_{30}C_{3k},\ \forall k \neq 0, \tag{69}$$

$$Q_1 = W_{01}C_{11} + W_{02}C_{12} + W_{03}C_{13}, \tag{70}$$

$$Q_2 = W_{01}C_{21} + W_{02}C_{22} + W_{03}C_{23}, \tag{71}$$

$$Q_3 = W_{01}C_{31} + W_{02}C_{32} + W_{03}C_{33}, \tag{72}$$

from Eqs. (18) and (19), with normalized probabilities given by $P_i = Q_i/\sum_j Q_j$.

Equation (48) gives the velocity and the dispersion is given by Eq. (49) with the C_{ij} as given in Eqs. (21)–(29).

A. Single chain

A toy 4-state single chain model with no backwards steps is shown in Figure 9 and has

$$Q_0^S = u_3^S C_{31}^S, \tag{73}$$

$$Q_1^S = u_0^S C_{11}^S, \tag{74}$$

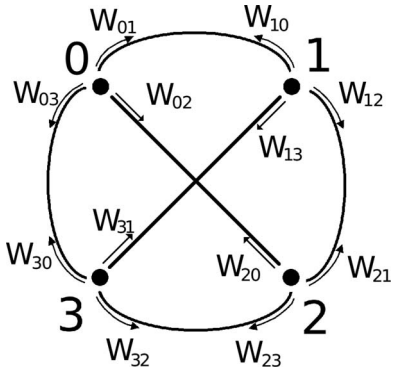

FIG. 8. A general 4-state system.

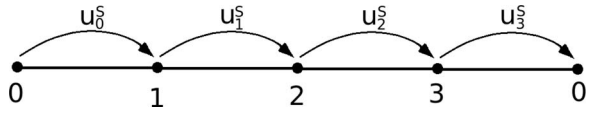

FIG. 9. A toy single chain model with no backwards stepping.

$$Q_2^S = u_0^S C_{21}^S, \tag{75}$$

$$Q_3^S = u_0^S C_{31}^S, \tag{76}$$

with $P_i^S = N^S Q_i^S$ and $N^S = 1/\sum_j Q_j^S$. Equation (48) gives the velocity

$$v^S = d u_3^S P_3^S. \tag{77}$$

The dispersion is given by Eq. (49)

$$D^S = \frac{d^2}{2} u_{n-1}^S P_{n-1}^S - \frac{dv^S}{Q_0^S} u_{n-1}^S \sum_{j\neq 0} P_j^S C_{n-1,j}^S + \frac{v^S v^S}{Q_0^S} \sum_i \sum_{j\neq 0} P_j^S C_{ij}^S, \tag{78}$$

with

$$C_{11}^S = u_2^S u_3^S, \tag{79}$$

$$C_{12}^S = 0, \tag{80}$$

$$C_{13}^S = 0, \tag{81}$$

$$C_{21}^S = u_1^S u_3^S, \tag{82}$$

$$C_{22}^S = u_1^S u_3^S, \tag{83}$$

$$C_{23}^S = 0, \tag{84}$$

$$C_{31}^S = u_1^S u_2^S, \tag{85}$$

$$C_{32}^S = u_1^S u_2^S, \tag{86}$$

$$C_{33}^S = u_1^S u_2^S. \tag{87}$$

B. Divided pathway with a jump

The toy 4-state divided pathway model with a jump and no backwards steps is shown in Figure 10. When $u_3^D = 0$, this is a divided pathway model. For simplicity, rates are chosen so that $u_{12} = u_{20}$ and $u_{13} = u_{30}$. For both of these systems,

$$Q_0^D = u_1^D C_{21}^D + u_2^D C_{31}^D, \tag{88}$$

$$Q_1^D = u_0^D C_{11}^D, \tag{89}$$

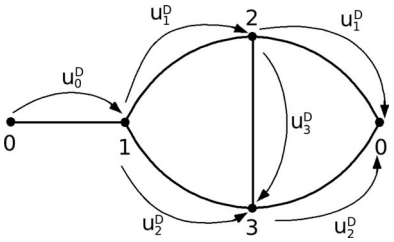

FIG. 10. A toy divided pathway with a jump model with no backwards stepping. When $u_3^D = 0$, this is a divided pathway model.

$$Q_2^D = u_0^D C_{21}^D, \tag{90}$$

$$Q_3^D = u_0^D C_{31}^D, \tag{91}$$

with $P_i^D = N^D Q_i^D$ and $N^D = 1/\sum_j Q_j^D$. Equation (48) gives the velocity

$$v^D = d\left[u_1^D P_2^D + u_2^D P_3^D\right]. \tag{92}$$

The dispersion is given by Eq. (49)

$$D^D = \frac{d^2}{2}\left(u_1^D P_2^D + u_2^D P_3^D\right) - \frac{dv^D}{Q_0^D} \sum_{j\neq 0} P_j^D \left(u_1^D C_{1j}^D + u_2^D C_{2j}^D\right) + \frac{v^D v^D}{Q_0^D} \sum_i \sum_{j\neq 0} P_j^D C_{ij}^D, \tag{93}$$

The randomness ratio is then expressed as

$$r^D = \frac{2D^D}{dv^D}. \tag{94}$$

For the divided pathway with a jump:

$$C_{11}^D = u_1^D u_2^D + u_2^D u_3^D, \tag{95}$$

$$C_{12}^D = 0, \tag{96}$$

$$C_{13}^D = 0, \tag{97}$$

$$C_{21}^D = u_1^D u_2^D, \tag{98}$$

$$C_{22}^D = u_2^D \left(u_1^D + u_2^D\right), \tag{99}$$

$$C_{23}^D = 0, \tag{100}$$

$$C_{31}^D = u_2^D \left(u_1^D + u_3^D\right) + u_1^D u_3^D, \tag{101}$$

$$C_{32}^D = u_3^D \left(u_1^D + u_2^D\right), \tag{102}$$

$$C_{33}^D = \left(u_1^D + u_3^D\right)\left(u_1^D + u_2^D\right), \tag{103}$$

with the divided pathway result recovered when $u_3^D = 0$. It should be noted that $C_{ij}^D = 0$ when all rate paths from j to i pass through 0; in this architecture, these rate paths are 2 to 1, 3 to 1, and 3 to 2.

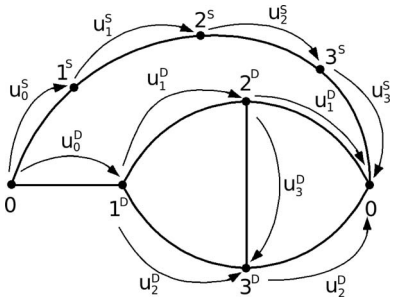

FIG. 11. A divided pathway with a jump and a single chain model with no backwards stepping.

C. Combining model structures

A divided pathway with a jump and a single chain (and no backwards stepping) is shown in Figure 11. As described in Sec. IV A, we can deduce the unnormalized probabilities to be

$$Q_0 = Q_0^S Q_0^D, \tag{104}$$

$$Q_{1^S} = Q_0^D Q_1^S, \tag{105}$$

$$Q_{2^S} = Q_0^D Q_2^S, \tag{106}$$

$$Q_{3^S} = Q_0^D Q_3^S, \tag{107}$$

$$Q_{1^D} = Q_0^S Q_1^D, \tag{108}$$

$$Q_{2^D} = Q_0^S Q_2^D, \tag{109}$$

$$Q_{3^D} = Q_0^S Q_3^D, \tag{110}$$

with $P_i = NQ_i$ with $N = 1/\sum_j Q_j$. Equation (48) gives the velocity

$$v = N\left[\frac{v^S}{N^S} + \frac{v^D}{N^D}\right]. \tag{111}$$

The dispersion is given by Eq. (49)

$$D = \frac{d^2}{2}\left(u_3^S P_{3^S} + u_2^D P_{2^S} + u_3^D P_{3^S}\right) - \frac{dv}{Q_0}\sum_{j\neq 0} P_j\left(u_3^S C_{3^S j} + u_2^D C_{2^S j} + u_3^D C_{3^S j}\right) + \frac{v^2}{Q_0}\sum_{j\neq 0} P_j \sum_i C_{ij}, \tag{112}$$

with

$$C_{ij} = C_{ij}^D Q_0^S,\ i, j \in \text{DJ section}, \tag{113}$$

$$C_{ij} = C_{ij}^S Q_0^D,\ i, j \in \text{S section}, \tag{114}$$

$$C_{ij} = 0,\ \text{otherwise}, \tag{115}$$

where *DJ* represents the divided pathway with a jump and *S* represents the single chain.

VII. DISCUSSION

We have presented a method to calculate measurable quantities of molecular motors described by discrete stochastic models. The approach is based on the work of Derrida[12] that gives dynamic properties for a general *n*-state single chain model. Derrida's work has in the past been extended to other system structures[13,14,16] and in each case the velocity and dispersion were rederived. However, our method—extending the work by King and Altman[17]—gives explicit expressions for the completely general steady-state probabilities, velocities, and dispersions in terms of variables C_{ij} that depend on the system structure.

The expressions for the average velocity and dispersion were derived without any constraint on the distance between physical stepping sites. It was shown that the resulting equations were independent of the substep size d_i (assuming that the transition rates are also) and only depend on the total step size d. While apparent for systems without detachment, this was not obvious *a priori* for systems with unequal detachments between between successive substeps. The renormalization procedure[16] maps a system with detachment to one without by scaling reaction rates and steady-state probabilities. Thus our results still hold with rescaled rates and probabilities and so the velocity and dispersion are independent of the substep sizes d_i regardless of detachment rates.

We have shown that the C_{ij} are derivable in a simple graphical manner from the structure of a proposed system and have written them down explicitly for several example structures. Our approach gives general *n*-state system expressions for multiple system structures. Modifications of the generalized structure and their effect on the C_{ij} have been explored. Results for a simple 4-state system demonstrate the simplicity of the calculations relative to other methods. Results for two separate 4-state systems can be combined in a simple manner to give the dispersion and velocity of an 8-state system.

We have given generalized results for the single chain, parallel pathway, and divided pathway systems and have used a technique for combining structures coupled at the reference state to derive the novel divided pathway with a chain results.

Alternative methods of calculating the dynamic properties exist. Tsygankov *et al.*[9] provide a flux-balance method to calculate the velocities for any system and Chemla *et al.*[11] use matrix methods to calculate the velocities and dispersions from a given system. However none can provide results for general model structures or solutions that can be interpreted in such a simple graphical way.

Our methods provide powerful theoretical tools for investigating how the underlying transition rates of a molecular motor affect its dynamic properties. The dynamic properties of smaller models can now be calculated simply. Large and highly complex models can be classified by their structure and our methods can give the steady-state probabilities and dynamic properties readily without the need to perform computationally expensive calculations.

This work can aid in the investigation of many biochemical systems. For example, the stepping mechanism of the linearly processive motor protein myosin-V is not understood in detail.[2,3,5,8,20–30] Our results allow the calculation of dynamic

properties for different postulated system architectures, thus can help to distinguish between them on the basis of the fit to experimental data.

ACKNOWLEDGMENTS

This work was supported by the Engineering and Physical Sciences Research Council Grant No. EP/P505135/1.

APPENDIX A: CALCULATION OF DYNAMIC QUANTITIES

In this appendix, we show how to calculate the expressions for the velocity (Eq. (14)) and dispersion (Eq. (16)).

1. Velocity

Multiplying Eq. (11) by s, summing over s and defining $X_i = \sum_s s p_{i,s}$ gives

$$\frac{dX_i}{dt} = \sum_{j=i+1}^{n-1}\left(w_{ji}X_j + u_{ji}\sum_s (s+1)p_{j,s}\right) + \sum_{j=0}^{i-1}\left(w_{ji}\sum_s (s-1)p_{j,s} + u_{ji}X_j\right) - \sum_j (u_{ij}+w_{ij})X_i, \tag{A1}$$

and so

$$\frac{dX_i}{dt} = \sum_j (W_{ji}X_j - W_{ij}X_i) + \sum_{j=i+1}^{n-1} u_{ji}P_j - \sum_{j=0}^{i-1} w_{ji}P_j, \tag{A2}$$

recognizing that $W_{ij} = u_{ij} + w_{ij}$ when mapping the physical-space system onto the state-space one.

Therefore,

$$\sum_i \frac{dX_i}{dt} = \sum_i \sum_{j=0}^{i-1}(u_{ji} - w_{ji})P_j. \tag{A3}$$

Substituting this into Eq. (13) gives the expression for the velocity in Eq. (14).

2. Dispersion

The the dispersion is given by

$$D \equiv \frac{1}{2}\lim_{t\to\infty}\frac{d}{dt}(\langle x^2\rangle - \langle x\rangle^2). \tag{A4}$$

The mean squared displacement is given by

$$\langle x^2\rangle = \sum_s\sum_i p_{i,s}(ds + d_i)^2, = \sum_i \left(d^2\alpha_i + 2dd_iX_i + d_i^2P_i\right), \tag{A5}$$

where $\alpha_i = \sum_s s^2 p_i$. Therefore, using Eq. (13), we have

$$D = \sum_i \left[d^2\frac{d\alpha_i}{dt} + 2dd_i\frac{dX_i}{dt} - 2vdX_i - 2vd_iP_i\right] \tag{A6}$$

at steady state.

Similarly to Eq. (A2), we have

$$\frac{d\alpha_i}{dt} = \sum_j (W_{ji}\alpha_j - W_{ij}\alpha_i) + \sum_{j=0}^{i-1}[w_{ji}(P_j - 2X_j) + u_{ji}(P_j + 2X_j)] \tag{A7}$$

and hence

$$D = d^2\sum_i\sum_{j=0}^{i-1}[w_{ji}(P_j - 2X_j) + u_{ji}(P_j + 2X_j)] + 2\sum_i\left(dd_i\frac{dX_i}{dt} - vdX_i - vd_iP_i\right). \tag{A8}$$

Therefore to calculate the dispersion, we must first find the X_i.

Assuming a constant velocity solution $X_i = g_i t + h_i$, the balance of constant terms in Eq. (A2) gives solutions for the h_i in terms of the g_i. Constant terms also sum to give a normalization condition for the g_i:

$$\sum_i g_i = \sum_i\sum_{j=0}^{i-1}(u_{ji} - w_{ji})P_j. \tag{A9}$$

Linear with time terms in Eq. (A2) are the same as the governing Eq. (3) for the steady-state probabilities P_i. Therefore, $g_i \propto P_i$ and normalizing gives

$$g_i = P_i\sum_k\sum_{j=0}^{k-1}(u_{jk} - w_{jk})P_j. \tag{A10}$$

Using $g_i = P_i v/d$, it can be seen that

$$D = \frac{d^2}{2}\lim_{t\to\infty}\sum_i\sum_{j=0}^{i-1}[w_{ji}(P_j - 2X_j) + u_{ji}(P_j + 2X_j)] - d\lim_{t\to\infty}v\sum_i X_i, = d^2\sum_i\lim_{t\to\infty}\left(\sum_{j=0}^{i-1}[u_{ji}X_j - w_{ji}X_j] - \frac{v}{d}X_i\right) + d^2\sum_i\frac{1}{2}\sum_{j=0}^{i-1}[u_{ji}P_j + w_{ji}P_j]. \tag{A11}$$

Only the h_i remain to be determined in order to calculate X_i.

Writing $h_i = A_ih_0 + B_i$, the constant terms in Eq. (A2) give

$$g_i = \sum_j[W_{ji}(A_jh_0 + B_j) - W_{ij}(A_ih_0 + B_i)] + \sum_{j=i+1}^{n-1}u_{ji}P_j - \sum_{j=0}^{i-1}w_{ji}P_j, \tag{A12}$$

matching up powers of h_0 give two equations

$$0 = \sum_j [W_{ji}A_j - W_{ij}A_i], \tag{A13}$$

$$g_i = \sum_j [W_{ji}B_j - W_{ij}B_i] + \sum_{j=i+1}^{n-1} u_{ji}P_j - \sum_{j=0}^{i-1} w_{ji}P_j. \tag{A14}$$

Equation (A13) gives $A_i = P_i/P_0$ and $B_0 = 0$. Therefore, the dispersion in Eq. (A11) can be written in terms of the B_i,

$$\begin{aligned} D &= d^2 \sum_i \left(\sum_{j=0}^{i-1} [u_{ji}B_j - w_{ji}B_j] - \frac{v}{d} B_i \right) \\ &\quad + d^2 \frac{h_0}{P_0} \sum_i \left(\sum_{j=0}^{i-1} [u_{ji}P_j - w_{ji}P_j] - \frac{v}{d} P_i \right) \\ &\quad + d^2 \sum_i \frac{1}{2} \sum_{j=0}^{i-1} [u_{ji}P_j + w_{ji}P_j], \\ &= d^2 \sum_i \sum_{j=0}^{i-1} \left[u_{ji}B_j - w_{ji}B_j + \frac{1}{2}(u_{ji}P_j + w_{ji}P_j) \right] \\ &\quad - dv \sum_i B_i, \end{aligned} \tag{A15}$$

with the only undetermined quantities being B_i. It should be noted that this expression is the steady-state dispersion and thus independent of the initial condition.

Equation (A14) can be written

$$G_i = \sum_j W_{ji}B_j - \sum_j W_{ij}B_i, \tag{A16}$$

with a suitably chosen G_i. Therefore, $\sum_j G_j = 0$.

This can be written

$$\mathbf{G} = \mathbb{M}\mathbf{B} \tag{A17}$$

with $B_0 = 0$. $\mathbb{M}$ is a singular matrix as $\mathbb{M}\mathbf{P} = 0$ with $\sum_i P_i = 1$. Equation (A17) cannot be solved explicitly for an arbitrarily sized system by standard matrix methods.

We have $B_0 = 0$ and $\sum_j G_j = 0$ and choose ansatz

$$B_i = -\frac{\sum_{j\neq 0} G_j C_{ij}}{Q_0}, \; i \neq 0, \tag{A18}$$

where again Q_0 is the sum over all configurations of state 0 with the C_{ij} chosen so that

$$C_{ij} = \frac{Z_{ij}}{W_{0j}}, \; i \neq 0, \tag{A19}$$

where the Z_{ij} are the sum over all configurations of state i that include a rate W_{0j}.

Equation (A16) and (A18) require that the C_{ij} satisfy

$$G_i Q_0 = \sum_k G_k \left[\sum_j W_{ij}C_{ik} - \sum_j W_{ji}C_{jk} \right] \tag{A20}$$

where again Q_0 is a sum over all configurations of state 0. Note that the $k = 0$ term in the summation is 0 by definition of the C_{ij}.

For $i = 0$, we have

$$\begin{aligned} &\sum_k G_k \left[\sum_j W_{0j}C_{0k} - \sum_j W_{j0}C_{jk} \right] \\ &= -\sum_{k\neq 0} G_k \sum_j W_{j0}C_{jk}, \\ &= -\sum_{k\neq 0} G_k Q_0, \text{ from Eq. (B2)} \\ &= G_0 Q_0, \end{aligned} \tag{A21}$$

since $\sum_j G_j = 0$. For $i \neq 0$, we have

$$\begin{aligned} &\sum_k G_k \sum_j [W_{ij}C_{ik} - W_{ji}C_{jk}] \\ &= G_i \sum_j [W_{ij}C_{ii} - W_{ji}C_{ji}] \\ &\quad + \sum_{k\neq i} G_k \sum_j [W_{ij}C_{ik} - W_{ji}C_{jk}], \\ &= \frac{G_i}{W_{0i}} \sum_j [W_{ij}Z_{ii} - W_{ji}Z_{ji}] \\ &\quad + \sum_{k\neq i} \frac{G_k}{W_{0k}} \sum_j [W_{ij}Z_{ik} - W_{ji}Z_{jk}], \\ &= G_i Q_0, \end{aligned} \tag{A22}$$

using Eqs. (B7) and (B14). Thus the choice of the C_{ij} in Eq. (A19) satisfies Eq. (A20) and therefore the ansatz (A18) is correct.

The dispersion in terms of the C_{ij} is therefore

$$\begin{aligned} D &= d^2 \sum_i \sum_{j=0}^{i-1} \frac{1}{2}(u_{ji} + w_{ji})P_j \\ &\quad + \frac{d^2}{Q_0} \sum_i \sum_{j=0}^{i-1} [w_{ji} - u_{ji}] \sum_{k\neq 0} G_k C_{jk} \\ &\quad + \frac{dv}{Q_0} \sum_i \sum_{j\neq 0} G_j C_{ij}, \end{aligned} \tag{A23}$$

with

$$G_i = \frac{v}{d} P_i - \sum_{j=i+1}^{n-1} u_{ji}P_j + \sum_{j=0}^{i-1} w_{ji}P_j, \tag{A24}$$

as given in Eqs. (16) and (17).

The C_{ij} are the sums over all configurations of state $i \neq 0$ given a rate from 0 to j and dividing through by that rate. Only these need to be calculated to give the state-occupancy probabilities, the velocity (Eq. (14)) and the dispersion (Eq. (16)). Systems with high degrees of symmetry can greatly simplify the calculation of the C_{ij}.

APPENDIX B: CONFIGURATIONAL METHODS

We shall define Z_{ik} and Z^*_{ik} to be the sum over all configurations and configurations* respectively of state i that include a W_{0k}. Thus we have $Z_{i0} = Z_{0i} = Z^*_{i0} = Z^*_{0i} = 0$ for all i. In the example shown in Figure 1, $Z_{1k} = 0$ for $k \neq 1, 3$, $Z_{13} = w_0 w_3 w_2$ and $Z_{11} = u_0(u_2 u_3 + w_2 u_3 + w_2 w_3)$.

It can be deduced that the sum over all configurations of $i \neq 0$ satisfies

$$Q_i = \sum_{k \neq 0} Z_{ik}, \; i \neq 0. \tag{B1}$$

Then Q_0 satisfies

$$Q_0 = \sum_{j \neq 0} W_{j0} \frac{Z_{jk}}{W_{0k}}, \; \forall k \neq 0 \tag{B2}$$

as required for the derivation of Eq. (A21). This can be seen by considering each element of the summation in turn. Each element j is the sum over all possible configurations of 0 containing the rate W_{j0}. Summing over all j must therefore give the sum over all configurations of 0, Q_0.

Denote a rate path from a to b that does not pass through state 0 by $W_{a \to b}$ and from a to b that does pass though state 0 by $W_{a \to 0 \to b}$. The corresponding closed rate paths are $\bar{W}_{a \to b}$ and $\bar{W}_{a \to 0 \to b}$, respectively.

A configuration* contains exactly one closed rate path by definition. Denote the sum over all configurations* of j that include a rate W_{0i} given that each term contains a $\bar{W}_{j \to i}$ by $Z^*_{ji}(\bar{W}_{j \to i})$ and similarly for $Z^*_{ji}(\bar{W}_{j \to 0 \to i})$ and $Z^*_{ji}(\bar{W}_{j \to i}$ or $\bar{W}_{j \to 0 \to i})$.

By definition, Z_{ii} is the sum over all configurations of i that include a non-zero rate W_{0i}. Therefore for any $j \neq 0, i$, each term contains either a $W_{j \to i}$, or a $W_{j \to 0 \to i}$. $Z_{ii} \sum_j W_{ij}$ is therefore the sum over all configurations* of i that include a non-zero rate W_{0i} given that each term contains either a $\bar{W}_{i \to i}$ or a $\bar{W}_{i \to 0 \to i}$ (Figure 12) and so

$$\begin{aligned} \sum_j W_{ij} Z_{ii} &= Z^*_{ii}(\bar{W}_{i \to i} \text{ or } \bar{W}_{i \to 0 \to i}), \\ &= Z^*_{ii}(\bar{W}_{i \to i}) + Z^*_{ii}(\bar{W}_{i \to 0 \to i}), \end{aligned} \tag{B3}$$

as a configuration* contains exactly one closed rate path.

Z_{ji} is the sum over all configurations of j that include a non-zero rate W_{0i}. Therefore, as configurations cannot contain closed rate paths, each term must contain a $W_{i \to j}$. Thus, each term in $\sum_j W_{ji} Z_{ji}$ must contain a $\bar{W}_{i \to i}$ and so $\sum_j W_{ji} Z_{ji}$ is the sum over all configurations* of i that include a non-zero rate W_{0i} given that each term contains a $\bar{W}_{i \to i}$ and so

$$Z^*_{ii}(\bar{W}_{i \to i}) = \sum_j W_{ji} Z_{ji}. \tag{B4}$$

Each term in $\sum_j W_{j0} Z_{ji}$ must contain a $\bar{W}_{i \to 0 \to i}$ and so $\sum_j W_{j0} Z_{ji}$ is the sum over all configurations* of i that include a non-zero rate W_{0i} given that each term contains a $\bar{W}_{i \to 0 \to i}$ and so

$$Z^*_{ii}(\bar{W}_{i \to 0 \to i}) = \sum_j W_{j0} Z_{ji}. \tag{B5}$$

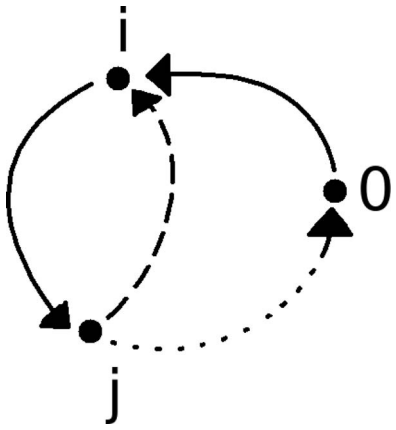

FIG. 12. The two possible classes of closed rate paths within the sum over all configurations* of i that include a rate W_{0i}, Z^*_{ii}. Each element of Z^*_{ii} either has a closed rate path from i to i through 0 ($\bar{W}_{i \to 0 \to i}$, dotted) or not through zero ($\bar{W}_{i \to i}$, dashed).

Using the above and then Eq. (B2),

$$\begin{aligned} \sum_j W_{ij} Z_{ii} &= \sum_j (W_{j0} + W_{ji}) Z_{ji}, \\ &= W_{0i} Q_0 + \sum_j W_{ji} Z_{ji}. \end{aligned} \tag{B6}$$

Therefore,

$$\sum_j [W_{ij} Z_{ii} - W_{ji} Z_{ji}] = W_{0i} Q_0, \tag{B7}$$

as required for the derivation of Eq. (A22).

Using the same argument used to show relation (7), it can be seen that

$$Z^*_{ik} = \sum_{j \neq 0,k} W_{ji} Z_{jk}, \tag{B8}$$

$$Z^*_{ik} = Z_{ik} \sum_{j \neq 0,k} W_{ij}, \tag{B9}$$

for $i \neq 0, k$. Therefore,

$$\sum_{j \neq 0,k} [W_{ij} Z_{ik} - W_{ji} Z_{jk}] = 0, \; \forall i \neq 0, k. \tag{B10}$$

Using the relations between Q_i and Z_{ij} in Eqs. (B1) and (B2), Eq. (7) gives

$$\begin{aligned} &W_{i0} \sum_{j \neq 0} Z_{ij} - W_{0i} Q_0 \\ &\quad + \sum_{j \neq 0} \sum_{k \neq 0} [W_{ij} Z_{ik} - W_{ji} Z_{jk}] = 0, \end{aligned} \tag{B11}$$

for all $i \neq 0$. Equations (B7) and (B10) then simplify this to

$$\sum_{j \neq 0} [(W_{i0} + W_{ij}) Z_{ij} - W_{ji} Z_{jj}] = 0. \tag{B12}$$

Each component of the sum over j contains a unique element W_{0j} and thus

$$(W_{i0} + W_{ij})Z_{ij} - W_{ji}Z_{jj} = 0, \tag{B13}$$

for all $i, j \neq 0$.

Therefore using Eqs. (B10) and (B13),

$$\sum_j [W_{ij}Z_{ik} - W_{ji}Z_{jk}] = 0, \ \forall i \neq 0, k, \tag{B14}$$

as required for the derivation of Eq. (A22).

APPENDIX C: DIVIDED PATHWAY C_{ij}

The divided pathway[14] (see Figure 5(c)) has two coupling states and three branches. For a given C_{ij}, states i and j can represent a coupling state and not just branch states. There are 14 different expressions for the C_{ij}, however it is sufficient to write down six and describe the transformations needed to produce the others.

The first five are given by

$$\begin{aligned} C_{ij} = u_j \Lambda_{j+1}^{i-1} \Xi_1^{j-1} \big[& w_m \Pi_{i+1}^{m-1\ k} \Xi_{m+1}^{n^k-1\ k'} \Xi_{m+1}^{n^{k'}-1} \\ & + \Xi_{i+1}^{m-1} \big(u_m^{k\ k} \Lambda_{m+1}^{n^k-1\ k'} \Xi_{m+1}^{n^{k'}-1} \\ & + u_m^{k'\ k'} \Lambda_{m+1}^{n^{k'}-1\ k} \Xi_{m+1}^{n^k-1} \big) \big], \end{aligned} \tag{C1}$$

for $0 < j < i < m$,

$$C_{ij} = u_j \Lambda_{j+1}^{i-1} \Xi_1^{j-1\ k} \Xi_{i+1}^{n^k-1\ k'} \Xi_{i+1}^{n^{k'}-1}, \tag{C2}$$

for $0 < j < i = m$,

$$C_{ij} = u_m^k \Xi_1^{j-1} \Lambda_j^{m-1\ k} \Lambda_{m-1}^{i-1\ k} \Xi_{i+1}^{n^k-1\ k'} \Xi_{m+1}^{n^{k'}-1}, \tag{C3}$$

for $0 < j < m < i$ with i on branch k,

$$C_{ij} = u_j^{k\ k} \Lambda_{j+1}^{i-1} \Xi_1^{j-1\ k} \Xi_{i+1}^{n^k-1\ k'} \Xi_{j+1}^{n^{k'}-1}, \tag{C4}$$

for $0 < j = m < i$ and

$$\begin{aligned} C_{ij} = u_j^{k\ k} \Lambda_{j+1}^{i-1\ k} \Xi_{i+1}^{n^k-1} & \\ \times \big[& w_m{}^k \Xi_{m+1}^{j-1} \Pi_1^{m-1\ k'} \Xi_{m+1}^{n^{k'}-1} \\ & + u_m^{k\ k} \Lambda_{m+1}^{j-1} \Xi_1^{m-1\ k'} \Xi_{m+1}^{n^{k'}-1} \\ & + u_m^{k'\ k'} \Lambda_{m+1}^{n^{k'}-1} \Xi_1^{m-1\ k} \Xi_{m+1}^{j-1} \big], \end{aligned} \tag{C5}$$

for $0 < m < j < i$.

The next five terms $0 < i \leq j < m$, $0 < i \leq j = m$, $0 < i < m < j$, $0 < i = m < j$, $0 < m < i \leq j$ are given by applying four transformations. The transformations given by relations (37), (38), and (39) and an additional $u_m^k \leftrightarrow w_m$ transformation to take into account the extra coupling state.

The cases where $0 < j = i < m$, $0 < j = i = m$, and $0 < m < j = i$ are given by Eqs. (C1), (C2), and (C5), respectively, with $u_j^{k\ k} \Lambda_{j+1}^{i-1} \to 1$.

The last term for $m < i, j$ with i on branch k and j on branch k' is

$$C_{ij} = u_m^k u_j^{k'\ k'} \Pi_{m+1}^{j-1\ k} \Lambda_{m+1}^{i-1} \Xi_1^{m-1\ k} \Xi_{i+1}^{n^k-1\ k'} \Xi_{j+1}^{n^{k'}-1}. \tag{C6}$$

[1] M. Rief, R. S. Rock, A. D. Mehta, M. S. Mooseker, R. E. Cheney, and J. A. Spudich, "Myosin-v stepping kinetics: A molecular model for processivity," Proc. Natl. Acad. Sci. U.S.A. **97**(17), 9482–9486 (2000).

[2] C. Veigel, F. Wang, M. L. Bartoo, J. R. Sellers, and J. E. Molloy, "The gated gait of the processive molecular motor, myosin-v," Nat. Cell Biol. **4**, 59–65 (2002).

[3] J. E. Baker, E. B. Krementsova, G. G. Kennedy, A. Armstrong, K. M. Trybus, and D. M. Warshaw, "Myosin-v processivity: Multiple kinetic pathways for head-to-head coordination," Proc. Natl. Acad. Sci. U.S.A. **101**(15), 5542–5546 (2004).

[4] S. S. Rosenfeld and H. L. Sweeney, "A model of myosin-v processivity," J. Biol. Chem. **279**(38), 40100–40111 (2004).

[5] S. Uemura, H. Higuchi, A. O. Olivares, E. M. De La Cruz, and S. Ishiwata1, "Mechanochemical coupling of two substeps in a single myosin-v motor," Nat. Struct. Mol. Biol. **11**(9), 877–883 (2004).

[6] J. C. M. Gebhardt, A. E.-M. Clemen, J. Jaud, and M. Rief, "Myosin-v is a mechanical ratchet," Proc. Natl. Acad. Sci. U.S.A. **103**(23), 8680–8685 (2006).

[7] K. I. Skau, R. B. Hoyle, and M. S. Turner, "A kinetic model describing the processivity of myosin-v," Biophys. J. **91**, 2475–2489 (2006).

[8] G. Cappello, P. Pierobon, C. Symonds, L. Busoni, J. C. M. Gebhard, M. Rief, and J. Prost, "Myosin-v stepping mechanism," Proc. Natl. Acad. Sci. U.S.A. **104**(39), 15238–15333 (2007).

[9] D. Tsygankov and M. E. Fisher, "Kinetic models for mechanoenzymes: Structural aspects under large loads," J. Chem. Phys. **128**, 015102 (2007).

[10] K. Svoboda, P. P. Mitra, and S. M. Block, "Fluctuation analysis of motor protein movement and single enzyme kinetics," Proc. Natl. Acad. Sci. U.S.A. **91**, 11782–11786 (1994).

[11] Y. R. Chemla, J. R. Moffitt, and C. Bustamante, "Exact solutions for kinetic models of macromolecular dynamics," J. Phys. Chem. B **112**(19), 6025–6044 (2008).

[12] B. Derrida, "Velocity and diffusion constant of a periodic one-dimensional hopping model," J. Stat. Phys. **31**(3), 433–450 (1983).

[13] A. B. Kolomeisky, "Exact results for parallel-chain kinetic models of biological transport," J. Chem. Phys. **115**(15), 7253–7259 (2001).

[14] R. K. Das and A. Kolomeisky, "Dynamic properties of molecular motors in the divided-pathway model," Phys. Chem. Chem. Phys. **11**, 4815–4820 (2009).

[15] E. B. Stukalin and A. B. Kolomeisky, "Transport of single molecules along the periodic parallel lattices with coupling," J. Chem. Phys. **124**, 204901 (2006).

[16] A. B. Kolomeisky and M. E. Fisher, "Periodic sequential kinetic models with jumping branching and deaths," Physica A **279**, 1–20 (2000).

[17] E. L. King and C. Altman, "A schematic method of deriving the rate laws of enzyme-catalyzed reactions," J. Phys. Chem. **60**, 1375–1378 (1956).

[18] T. L. Hill, *Free Energy Transduction and Biochemical Cycle Kinetics* (Springer-Verlag, New York, 1989).

[19] Y. Wu, Y. Q. Gao, and M. Karplus, "A kinetic model of coordinated myosin-v," Biochemistry **46**, 6318–6330 (2007).

[20] A. D. Mehta, R. S. Rock, M. Rief, J. A. Spudich, M. S. Mooseker, and R. E. Cheney, "Myosin-v is a processive, actin-based motor," Nature (London) **400**, 590–593 (1999).

[21] E. M. De La Cruz, A. L. Wells, S. S. Rosenfeld, E. M. Ostap, and H. L. Sweeney, "The kinetic mechanism of myosin-v," Proc. Natl. Acad. Sci. U.S.A. **96**(24), 13727–13731 (1999).

[22] E. M. De La Cruz, H. L. Sweeney, and E. M. Ostap, "Adp inhibition of myosin-v atpase activity," Biophys. J. **79**, 1524–1529 (2000).

[23] A. D. Mehta, "Myosin learns to walk," J. Cell. Sci. **114**, 1981–1998 (2001).

[24] C. M. Yengo, E. M. De La Cruz, D. Safer, M. Ostap, and H. L. Sweeney, "Kinetic characterization of the weak binding states of myosin-v," Biochemistry **41**, 8507–8517 (2002).

[25] J. N. Forkey, M. E. Quinlan, M. A. Shaw, J. E. T. Corrie, and Y. E. Goldman, "Three-dimensional structural dynamics of myosin-v by single-molecule fluorescence polarization," Nature (London) **422**, 399–404 (2003).

[26] G. E. Snyder, T. Sakamoto, J. A. Hammer, J. R. Sellers, and P. R. Selvin, "Nanometer localization of single green fluorescent proteins: Evidence that myosin-v walks hand-over-hand via telemark configuration," Biophys. J. **87**, 1776–1783 (2004).

[27]C. Veigel, S. Schmitz, F. Wang, and J. R. Sellers, "Load-dependent kinetics of myosin-v can explain its high processivity," Nat. Cell Biol. **7**(9), 861–869 (2005).
[28]S. Schmitz, J. Smith-Palmer, T. Sakamoto, J. R Sellers, and C. Veigel, "Walking mechanism of the intracellular cargo transporter myosin-v," J. Phys.: Condens. Matter **18**, S1943–S1956 (2006).
[29]E. Forgacs, S. Cartwright, T. Sakamoto, J. R. Sellers, J. E. T. Corrie, M. R. Webb, and H. D. White, "Kinetics of adp dissociation from the trail and lead heads of actomyosin-v following the power stroke," J. Biol. Chem. **283**(2), 766–773 (2008).
[30]A. Vilfan, "Five models for myosin-v," Front. Biosci. **14**, 2269–2284 (2005).