A Matrix Computational Approach to Kinesin Neck Linker Extension

John Hughes[a], William O. Hancock[b], John Fricks[a,*]

[a] *Department of Statistics, The Pennsylvania State University, University Park, PA 16802*

[b] *Department of Bioengineering, The Pennsylvania State University, University Park, PA 16802*

Abstract

Kinesin stepping requires both tethered diffusion of the free head and conformational changes driven by the chemical state of the motor. We present a numerical method using matrix representations of approximating Markov chains and renewal theory to compute important experimental quantities for models that include both tethered diffusion and chemical transitions. Explicitly modeling the tethered diffusion allows for exploration of the model under perturbation of the neck linker; comparisons are made between the computed models and from *in vitro* assays.

Keywords: molecular motors, kinesin, stochastic models, chemical kinetics, Brownian motion, renewal reward process, stochastic differential equations, flashing ratchet, semi-Markov process

1. Introduction

Within eukaryotic cells, small molecules such as sugars and amino acids move to where they are needed by the passive process of molecular diffusion. But diffusion is inadequate for the transport of larger cargos such as vesicles and organelles. A mitochondrion is approximately 10,000 times as large as a glucose molecule, for example. An object so large will not only be slow to

*Corresponding author. Tel +1-814-865-3235, Fax +1-814-863-7114

Email addresses: jph264@psu.edu (John Hughes), wohbio@engr.psu.edu (William O. Hancock), fricks@stat.psu.edu (John Fricks)

[1]

diffuse due to its mass but will also likely be impeded by collisions with other objects in the crowded cytoplasm. Consequently, cells have evolved active transport processes to supplement passive ones. An important class of agents of active transport are motor proteins that hydrolyze ATP to move along a cytoskeletal track. One such protein is kinesin, of which over 40 varieties have been identified in humans alone (Miki et al. (2001)).

A conventional kinesin molecular motor comprises two heads, a neck linker, a coiled-coil rod, and a cargo-binding tail (Hirokawa et al. (1989), Yang et al. (1989)). The motor "walks" along a microtubule using the two heads each joined to the coiled-coil by the neck linker domain and dragging a payload bound to the tail. Eventually the motor dissociates from the microtubule after having taken some random number of steps (usually in the hundreds). Each kinesin step is governed by a series of chemical reactions including ATP binding and hydrolysis, along with tethered diffusion, which occurs after the stepping head unbinds from the microtubule and before it rebinds to the microtubule. A step can be either forward or backward, but the motor's mechanochemical mechanism makes backward steps rare. Figure 1 shows one proposed model for the chemical steps that contribute to a physical step (Hancock and Howard (2003), Block (2007),Vale and Milligan (2000),Schief and Howard (2001),Cross (2004),Muthukrishnan et al. (2009)).

The neck linker domain, a fourteen residue polypeptide that connects the core motor domain to the coiled-coil dimerization domain, plays a key role in kinesin stepping. Docking of the neck linker to the bound head domain is thought to provide principal conformational change that drives kinesin stepping (Rice et al. (1999), Block (2007)). In addition, the neck linker of the free head acts as a tether that restricts the diffusional search of the free head for its next binding site. If the neck linker is too short, then the free head cannot reach its proper forward binding site, but if it is too long, then the head is able to easily bind to adjacent or rearward binding sites. The role of the neck linker in this tethered diffusion is currently a question of particular interest due to recent reports showing that altering the length of the neck linker significantly alters motor function (Hackney et al. (2003),Muthukrishnan et al. (2009),Yildiz et al. (2008),Shastry and Hancock (2010)). Furthermore, it has been shown that neck linker length varies considerably in diverse kinesins, leading to the hypothesis that this structural feature tunes motors for their specific intracellular tasks (Hariharan and Hancock (2009)). To better understand kinesin's chemomechanical mechanism and uncover specific mechanistic features of unconventional kinesins, quantitative models are needed

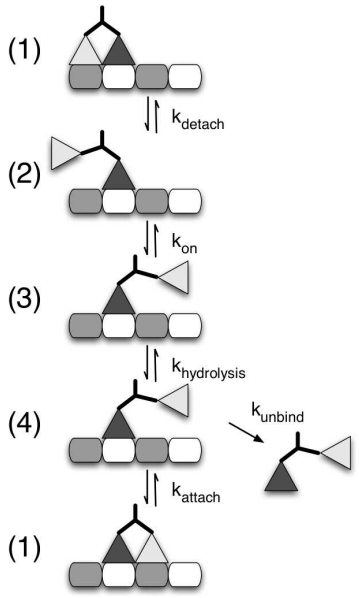


Figure 1: A proposed model for the chemical reactions of a kinesin step. Note that the rate k_{on} is linearly dependent on the concentration of ATP.

that incorporate both the chemical cycle as well as tethered diffusion of the free head during kinesin stepping.

This combination of chemical reactions and diffusion determines both the duration and direction of a step and also accounts for the dissociation of the motor from the microtubule. We will refer to the duration, or dwell time, of the ith step as τ_i and the direction and direction of the ith step as Z_i. In this paper, we will present computational techniques to determine the asymptotic behavior of $X(t) = L\sum_{i=1}^{N(t)} Z_i$, the approximate location of the motor at time t where $N(t)$ is the number of cycles completed by time t and L is the size of the step. We will define a mechanochemical cycle in more detail later, but essentially a cycle can be viewed as the time between having both heads bound until both heads are bound again.

While there have been a number of proposed models that include the combination of chemical and diffusive elements, these tend to rely on the spatially periodic structure of the microtubule in order to create a similarly spatial periodic model for the motor. (See reviews of molecular motor modeling Julicher et al. (1997), Kolomeisky and Fisher (2007), and Mogilner et al.

(2002)). An example of such a model is the flashing ratchet where the location of the molecular motor is described as a Brownian particle in one of several periodic potentials depending on the chemical state of the process. While this is an important and useful abstraction, it is difficult to include neck linker extension where the free motor head may move beyond a single spatial period several times before binding.

The framework presented here allows for relatively quick calculation of standard experimental quantities across a wide range of physical parameters, including the range of motion for the free motor head. While the current work focuses on kinesin, the method should be adaptable to other motors such as myosin 5, myosin 6, and dynein (Mallik et al. (2004), Rock et al. (2005), Sellers and Veigel (2006)). The matrix analytic framework is similar to the method presented by Wang, Peskin, and Elston, (WPE method) but by relying on a renewal-reward type formulation the current method can be more easily applied to alternative motor models which include a more detailed description of the motor sub-step dynamics (Wang et al. (2003)). Similar to the WPE method, the current presentation relies on a discretization of space rather than time; however, the approach here differs from the WPE method in that it does not rely on stationarity of the position of the free head within a cycle. The WPE method assumes that the motor can be described as moving continuously through a spatially periodic potential possibly switching between chemical states modeled by a Markov chain and takes advantage of the fact that the process modulo the period is stationary. This method was extended to include a cargo which is coupled to the motor in Fricks et al. (2006) and Xing et al. (2005). Both of these extensions rely on the fact that the distance from the motor to the cargo is also a stationary quantity.

The present work shifts the perspective from a strictly periodic spatial structure to a decomposition that relies on dividing time into suitably regularized stochastic quantities in order to handle the two alternating heads of kinesin. This is done by observing that as the motor walks down the microtubule, one head must stay bound and the other will move to a new binding site either returning to the binding site from which it detached or moving two binding sites away (hopping over the attached head). While it is more common for the rear head to detach first, enabling it to move forward and bind to the next free site, the front head can also detach and either return to that site or move to backward bind behind the currently bound head. A cycle then will shift from moving through a purely periodic potential to a recurrence time from the attachment of the free head until one head comes

unbound and then reattaches. When both heads are bound, the dynamics of the motor are essentially restarted and the motor does not have any memory as to whether it has taken a forward or backward step (or no step at all).

While time course simulation of stochastic processes can be more straightforward in implementation, they often require massive amounts of computational time and storage making sensitivity analysis especially difficult. In addition, since errors are accumulated through time with such methods, accurate computation of the distribution of the time of a step can be especially difficult. The renewal-reward approach has been applied to similar diffusive models such as in Lindner et al where these methods are used to analyze a Brownian particle in a periodic potential–one type of basic motor mode, and those methods were extended to handle kinetic switching of the periodic potential by Latorre et al. (Lindner et al. (2001), Latorre et al. (2007)) However, the method presented here allows for renewal events other than hitting times. Wang and Qian also introduced related methods through a study of semi-Markov processes and kinetic models for motors, and Das and Kolomeisky introduced methods to compute experimental quantities from similar models (Wang and Qian (2007), Das and Kolomeisky (2009)).

The main results and organization of this paper can be summarized as follows:

- A framework is presented for the analysis of processive molecular motors through the use of renewal-reward processes and extensions.
- Formulae are derived for standard asymptotic quantities within this framework.
- Quantities for within-step dynamics are computed using matrix analytic techniques.
- A model of kinesin which includes chemical and mechanical attributes of stepping is represented in this matrix analytic framework.
- Experimental data from neck linker extensions are accounted for using the numerical exploration of the model.

Section 2 presents a renewal-reward description of motor dynamics and includes an introduction to common quantities of interest in the molecular motor literature. In section 3, we show how to link a simple chemical kinetic model to the renewal-reward structure. In section 4, we present a model

that includes both chemical kinetics and spatial diffusion and show that an approximating Markov chain can fit into a matrix structure similar to the purely kinetic model. In addition, we show how the discrete space approximation can be used to explore the effect of neck linker extension in the context of particular neck linker models.

2. Renewal-Reward Structure

During a step, a motor has one bound head and one free head. We assume for the moment that the motor steps by taking only forward steps (at random times), and we denote the number of steps taken up to time t by $N(t)$. We assume that the motor goes through steps with times, τ_i, that are independent and identically distributed with mean μ_τ and variance σ_τ^2. Given these assumptions, $N(t)$ is an example of a renewal process, and the times between events are called renewals. Hence, the position of the motor (or, more precisely, the bound head) can be written as $X(t) = LN(t)$, where L is the length of a step.

An important quantity in the study of molecular motors is *asymptotic velocity* and is generally defined as

$$V_\infty = \lim_{t\to\infty} \frac{\mathbb{E}X(t)}{t}. \tag{1}$$

Given the above definition of $X(t)$, we have

$$V_\infty = \lim_{t\to\infty} \frac{\mathbb{E}LN(t)}{t}. \tag{2}$$

From the theory of renewal processes (Cox (1962)),

$$\mathbb{E}N(t) = \frac{t}{\mu_\tau} - \frac{1}{2} + o(1). \tag{3}$$

where $o(1)$ represents a quantity that converges to zero as $t \to \infty$ implying $V_\infty = L/\mu_\tau$. But the renewal theorem says something stronger, namely,

$$\lim_{t\to\infty} \frac{X(t)}{t} = \lim_{t\to\infty} \frac{LN(t)}{t} = \frac{L}{\mu_\tau} \tag{4}$$

with probability one, a useful fact in analyzing molecular motor data. If one can follow a *single* long path, the distance travelled divided by the travel time yields the velocity, i.e., multiple paths are not required.

Another important quantity in the study of motors is *effective diffusion* (Mogilner et al. (2002)) and can be defined as

$$D = \lim_{t\to\infty} \frac{\mathbb{V}X(t)}{2t}. \tag{5}$$

The theory of renewal processes also gives the variance of $N(t)$ as:

$$\mathbb{V}N(t) = \frac{\sigma_\tau^2}{\mu_\tau^3}t + \left(\frac{1}{12} + \frac{5\sigma_\tau^4}{4\mu_\tau^4} - \frac{2\eta_\tau}{\mu_\tau^3}\right) + o(1), \tag{6}$$

if we assume a finite third moment of the step times, η_τ. Thus, for the case of a forward-stepping motor,

$$D = \lim_{t\to\infty} \frac{\mathbb{V}LN(t)}{2t} = \frac{L^2\sigma_\tau^2}{2\mu_\tau^3}. \tag{7}$$

Now we can extend to the case of the motor taking both forward and backward steps. Let Z_i be the number of microtubule binding sites (forward or backward) in the ith step. Also, suppose that Z_i has mean μ_z and variance σ_z^2. Then

$$X(t) = L\sum_{i=1}^{N(t)} Z_i. \tag{8}$$

If we assume that the Z_i's are independent of one another and of $N(t)$, then $N(t)$ is again a renewal process, and $X(t)$ is an example of a renewal-reward process. Using the definitions for asymptotic velocity and effective diffusion above, we can calculate these quantities in a straightforward manner.

$$\begin{aligned}
\mathbb{E}X(t) &= \mathbb{E}\left[L\sum_{i=1}^{N(t)} Z_i\right] && (9)\\
&= L\mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{N(t)} Z_i \mid N(t)\right]\right] && (10)\\
&= L\mathbb{E}\left[N(t)\right]\mu_z && (11)\\
&\approx L\mu_z\frac{t}{\mu_\tau}, && (12)
\end{aligned}$$

and the asymptotic velocity is then

$$V_\infty = \frac{L\mu_z}{\mu_\tau} \tag{13}$$

for this case that includes backward steps.

We can calculate something similar for the effective diffusion. First, note that for any two random variables X and Y,

$$\mathbb{V}X = \mathbb{E}\left[\mathbb{V}\left[X \mid Y\right]\right] + \mathbb{V}\left[\mathbb{E}\left[X \mid Y\right]\right]. \tag{14}$$

Using this fact, we obtain

$$\begin{aligned}
\mathbb{V}\left[X(t)\right] &= \mathbb{V}\left[L\sum_{i=1}^{N(t)} Z_i\right] && (15)\\
&= L^2\mathbb{V}\left[\mathbb{E}\left[\sum_{i=1}^{N(t)} Z_i \mid N(t)\right]\right] + L^2\mathbb{E}\left[\mathbb{V}\left[\sum_{i=1}^{N(t)} Z_i \mid N(t)\right]\right] && (16)\\
&= L^2\mathbb{V}\left[N(t)\mu_z\right] + L^2\mathbb{E}\left[\sigma_z^2 N(t)\right] && (17)\\
&\approx L^2\mu_z^2\frac{\sigma_\tau^2}{\mu_\tau^3}t + L^2\sigma_z^2\frac{t}{\mu_\tau}. && (18)
\end{aligned}$$

Thus, the effective diffusion will be

$$D = \frac{L^2}{2}\left(\frac{\mu_z^2\sigma_\tau^2}{\mu_\tau^3} + \frac{\sigma_z^2}{\mu_\tau}\right) \tag{19}$$

for this more general case.

The calculations for asymptotic velocity and effective diffusion can also be used to formulate a functional central limit theorem which gives a clear interpretation for effective diffusion. Specifically, by scaling time by a factor of n and the step sizes by $n^{-1/2}$, then taking the limit as n gets large, we have

$$n^{-1/2}\left(X(nt) - V_\infty nt\right) \Rightarrow_n \sqrt{2D}\, B(t), \tag{20}$$

where $B(t)$ is a standard Brownian motion (Whitt (2002)). This result also implies a more conventional central limit theorem:

$$\frac{X(t) - V_\infty t}{\sqrt{2Dt}} \Rightarrow_t \mathcal{N}, \tag{21}$$

where $\mathcal{N}$ is a standard normal random variable. This can again be useful for data analysis as we can rewrite this expression as

$$\frac{X(t)}{t} \dot{\sim} V_\infty + \sqrt{\frac{2D}{t}}\mathcal{N}, \tag{22}$$

i.e., the observed velocity is approximately normally distributed.

If we can show that the time between steps is independent of the step direction, then by calculating the first two moments of the distributions of the times and step directions, we can easily calculate these asymptotic quantities of interest, effective diffusion and asymptotic velocity. The functional central limit theorem further confirms that, in some sense, these are the two quantities that matter since they are sufficient to define the limiting process.

A renewal reward framework is a natural model for the forward and backward stepping of a molecular motor. However, as we will see below, the assumption of independence between the step direction and time between steps may not be correct for a number of reasonable models. If we now assume that Z_i and τ_i are dependent, we can modify the law of large numbers and the functional central limit theorem for renewal-reward processes to handle this case.

We can continue to take as our definition of the process the expression (8). We simply note that the Z_i and τ_i are dependent on one another with covariance $\sigma_{Z,\tau}$. The law of large numbers continues to hold regardless of the dependence between the Z_i and $N(t)$. (See Billingsley (2008) for a proof.)

$$\lim_{t\to\infty} \frac{X(t)}{t} = \frac{L\mu_Z}{\mu_\tau} \tag{23}$$

We can also modify the functional central limit theorem to handle the dependency between time and jump direction by modifying the result for the renewal reward process. The result for the renewal reward process as presented in Whitt relies on a functional central limit theorem for the sums of Z_i and τ_i (Whitt (2002)). The modification to our situation yields the same limit theorem as in (20), but with an alternative definition of the effective diffusion,

$$2D = L^2\left(\frac{\sigma_Z^2}{\mu_\tau} + \frac{\mu_z^2\sigma_\tau^2}{\mu_\tau^3} - 2\frac{\mu_Z\sigma_{Z,\tau}}{\mu_\tau^2}\right).$$

The derivation is shown in the appendix. Note that when the correlation is set to zero we recover the definition for the renewal reward example.

One aspect that is missing from this framework, however, is that a motor has only a finite space over which to walk. In addition, a motor eventually dissociates from the microtubule. While the asymptotic quantities discussed above can be useful, we need to further consider *processivity* through the average run length of a motor, i.e., how far a motor travels, on average, before dissociation. This quantity is commonly used in experimental settings (Block et al. (1990), Vale et al. (1996)).

We will denote the number of steps prior to dissociation as N. It is natural to model N as a geometric random variable with probability of "success" equal to r, the probability of dissociating from the microtubule. We can then write the run length as

$$R = L \sum_{i=1}^{N} Z_i, \tag{24}$$

where Z_i is once again the direction of the ith step and L is the step length (typically 8 nm for kinesin). However, if the ith cycle ends in dissociation, $Z_i = 0$, which occurs with probability r. Using this model, the mean run length is

$$\mathbb{E}R = L\mathbb{E}N\mathbb{E}Z_i = L\frac{\mu_z}{r}. \tag{25}$$

3. An Illustrative Example–Kinetic Stepping Model

A common motor model is to assume that the motor must pass through a sequence of kinetic states before taking a physical step–either forward or backward as described in the introduction (Hancock and Howard (2003),Block (2007), Vale and Milligan (2000), Schief and Howard (2001), Cross (2004)). We could think, then, of a continuous time Markov chain that has transition rates that are periodic in space. We further assume that we cannot jump from one period to the next without going through a specific chemical step. For simplicity, we consider a four state motor model diagramed in Figure 1, where the motor can cycle through states one to four or in reverse.

In order to find the asymptotic quantities of interest introduced in the previous section, we need to define what constitutes a step. In the present context, we will define a step to be complete when both heads are bound after one has become unbound. In the context of this simplified pure chemical model, the motor will begin in chemical state one and end when the motor

returns to chemical state one. During this cycle, the motor has either moved forward one binding site, backward one binding site, or has returned to the same binding site. If we take Z_i to be the displacement of the leading head of the motor after one cycle, then the movements described correspond to one, negative one, and zero respectively. Note that there is some ambiguity in the definition of a cycle; we could have defined a cycle as the time after leaving chemical state one until arriving at chemical state one in a forward step or arriving at chemical state one in a backward step. The important feature is that once a cycle is complete, the evolution of the next cycle is independent of the past cycles.

This framework allows us to concentrate on the actions of a single step in which the motor starts in state one, and the cycle stops when both heads are again bound. If we can compute the moments of both the stopping time, τ, and the random variable associate with stepping, Z, including the cross moment between the two, then we can calculate the asymptotic quantities of interest and the processivity. In this local context then, we have a Markov chain, $Y(t)$, that is run until being absorbed into one of the states that corresponds to the end of a cycle. This local chain will then determine the distribution of τ and Z which in turn define the distribution for the position of the motor, $X(t)$.

We will assume that a physical step has occurred if the motor leaves state one and returns to state one. So, one way to describe the rate matrix of a single step is the following block form:

$$\mathbf{Q} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \tag{26}$$

where $\mathbf{A}$ represents the evolution within the cycle and $\mathbf{B}$ includes transitions to the absorbing states. (The $\mathbf{0}$ block matrices are filled in to make $\mathbf{Q}$ a square matrix.) Specifically, for the corresponding chemical cycle shown in

Figure 1, $\mathbf{A}$ will be defined as

$$\mathbf{A} = \begin{pmatrix} k_{1_+,1_+} & k_{1_+,2_+} & 0 & 0 & k_{1_+,4_-} & 0 & 0 \\ 0 & k_{2_+,2_+} & k_{2_+,3_+} & 0 & 0 & 0 & 0 \\ 0 & k_{3_+,2_+} & k_{3_+,3_+} & k_{3_+,4_+} & 0 & 0 & 0 \\ 0 & 0 & k_{4_+,3_+} & k_{4_+,4_+} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & k_{4_-,4_-} & k_{4_-,3_-} & 0 \\ 0 & 0 & 0 & 0 & k_{3_-,4_-} & k_{3_-,3_-} & k_{3_-,2_-} \\ 0 & 0 & 0 & 0 & 0 & k_{2_-,3_-} & k_{2_-,2_-} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{27}$$

and

$$\mathbf{B} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & k_{2_+,1_*} & 0 \\ 0 & 0 & 0 \\ k_{4_+,1_{++}} & 0 & 0 \\ 0 & k_{4_-,1_*} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & k_{2_-,1_-} \end{pmatrix}. \tag{28}$$

Note that $k_{i,i}$ will be the negative of the sum of the non-diagonal entries of the ith row of $\mathbf{A}$ and the entries of the ith row of matrix $\mathbf{B}$. The states 1_{++} and 1_- are used to denote chemical state 1 at the binding site forward and backward, respectively. The state 1_* corresponds to a return to chemical state one with the leading head returning to the same binding site as when the cycle was started.

As discussed in the previous section, important quantities of interest to calculate for this model are 1) the time until reaching the next cycle (or more specifically the moments from the distribution of the time) and 2) the distribution of the step direction. To compute these, we need to use the transition probability function for the local Markov chain, $\mathbf{P}_{i,j}(t)$. We can represent these functions in matrix form, and they must satisfy

$$\mathbf{P}'(t) = \mathbf{Q}\mathbf{P}(t), \tag{29}$$

where $\mathbf{P}_{i,j}(t) = \mathbb{P}(Y(t) = j \,|\, Y(0) = i)$ with the initial condition of $\mathbf{P}(0) = \mathbf{I}$. See Karlin and Taylor for details (Karlin and Taylor (1975)). The solution

to this equation is given as

$$\mathbf{P}(t) = e^{\mathbf{Q}t} = \sum_{i=0}^{\infty} \frac{t^i}{i!}\mathbf{Q}^i. \tag{30}$$

We will use the specific structure of $\mathbf{Q}$ to make this formula more useful. First, we assume that $\mathbf{A}$ is invertible. This is true if all of the states of $\mathbf{A}$ are transient and if all the states communicate with absorbing states (Neuts (1994)). Note that this decomposition and the exposition which follows is quite similar to that found in Colquhoun and Hawkes (1982) and Fredkin and Rice (1986) where they were used for applications in neuroscience.

We will take advantage of the particular structure of $\mathbf{Q}$ to simplify $\mathbf{P}(t)$. We note that

$$\mathbf{Q}^i = \begin{pmatrix} \mathbf{A}^i & \mathbf{A}^{i-1}\mathbf{B} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{A}^i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{31}$$

for $i \geq 1$. Hence, we can write

$$e^{\mathbf{Q}t} = \mathbf{I} + \sum_{i=1}^{\infty} \frac{t^i}{i!} \begin{pmatrix} \mathbf{A}^i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}. \tag{32}$$

Adding and subtracting an identity matrix yields

$$e^{\mathbf{Q}t} = \mathbf{I} + \left(\begin{pmatrix} \sum_{i=0}^{\infty} \frac{t^i}{i!}\mathbf{A}^i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} - \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \right) \begin{pmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \tag{33}$$

that gives

$$e^{\mathbf{Q}t} = \begin{pmatrix} e^{\mathbf{A}t} & e^{\mathbf{A}t}\mathbf{A}^{-1}\mathbf{B} - \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}. \tag{34}$$

Thus, if $\mathbf{A}$ is invertible, we need only find the matrix exponential of $\mathbf{A}$ and not $\mathbf{Q}$.

This formulation allows us to write down a distribution for the time to reach one of the absorbing states, i.e. τ in the previous discussion, when starting in chemical State 1 (from Figure 1). The cumulative distribution function to arrive in an absorbing state is then

$$\begin{aligned} F(t) &= \mathbb{P}(Y(t) = 1_- \text{ or } 1_{++} \text{ or } 1_*) && (35) \\ &= 1 - \mathbb{P}(Y(t) \neq 1_- \text{ or } 1_{++} \text{ or } 1_*) && (36) \\ &= 1 - \boldsymbol{a}' e^{\mathbf{A}t}\mathbf{1}, && (37) \end{aligned}$$

where $\boldsymbol{a}$ is the vector describing the probability of starting in one of the transient states (in this case, a one in the fourth position and zeros elsewhere) and $\mathbf{1}$ is a vector of ones. The distribution of the time to absorption in a finite state Markov chain is sometimes known as the phase distribution and has been well studied in the queueing literature (Asmussen (2003)). The density is given by

$$f(t) = -\boldsymbol{a}'\mathbf{A}e^{\mathbf{A}t}\mathbf{1}. \tag{38}$$

From this we can calculate the mean and variance of the recurrence time, μ_τ and σ_τ^2. For the mean, we have

$$\mu_\tau = -\boldsymbol{a}' \int_0^\infty t\mathbf{A}e^{\mathbf{A}t}dt\,\mathbf{1}. \tag{39}$$

Integration by parts (and assuming full rank of the submatrix) gives

$$\mu_\tau = -\boldsymbol{a}'te^{\mathbf{A}t}\mathbf{1}\Big|_0^\infty - \boldsymbol{a}' \int_0^\infty e^{\mathbf{A}t}dt\,\mathbf{1}. \tag{40}$$

Noting that for large time the matrix exponential must be approaching zero at an exponential rate, we have

$$\mu_\tau = -\boldsymbol{a}' \int_0^\infty e^{\mathbf{A}t}\mathbf{A}dt\mathbf{A}^{-1}\mathbf{1} = -\boldsymbol{a}'\mathbf{A}^{-1}\mathbf{1}. \tag{41}$$

A similar calculation, applying integration by parts twice, yields

$$\mathbb{E}\tau^2 = -\boldsymbol{a}' \int_0^\infty t^2\mathbf{A}e^{\mathbf{A}t}dt\,\mathbf{1} = 2\boldsymbol{a}'\mathbf{A}^{-2}\mathbf{1}. \tag{42}$$

And so the variance is

$$\sigma_\tau^2 = 2\boldsymbol{a}'\mathbf{A}^{-2}\mathbf{1} - \left(\boldsymbol{a}'\mathbf{A}^{-1}\mathbf{1}\right)^2. \tag{43}$$

We now have a straightforward way to calculate the mean and variance of the recurrence time. We now need the mean and variance of the step distribution. We can calculate the probability of arriving in one of the absorbing states at a time t and then take a limit in time to find the marginal distribution for Z. The joint probability of arriving at a particular absorbing state before time t can be written as

$$P(Z = z, \tau \leq t) = (\boldsymbol{a}|\mathbf{0})'e^{\mathbf{Q}t}(\mathbf{0}|\boldsymbol{c}) \tag{44}$$

where $\boldsymbol{a}$ is a vector of a length equal to the number of transient states with a one corresponding to chemical state one, the starting point of a cycle, and $\boldsymbol{c}$ is a vector with length equal to the number of absorbing states with a one in the position corresponding to the absorbing state, z. The zero vectors are concatenated to these vectors to ensure the correct dimensions for multiplication. We can then simplify the calculation:

$$P(Z = z, \tau \leq t) = (\boldsymbol{a}|\mathbf{0})' \begin{pmatrix} e^{\mathbf{A}t} & e^{\mathbf{A}t}\mathbf{A}^{-1}\mathbf{B} - \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} (\mathbf{0}|\boldsymbol{c}) \tag{45}$$

$$= \boldsymbol{a}' \left(e^{\mathbf{A}t}\mathbf{A}^{-1}\mathbf{B} - \mathbf{A}^{-1}\mathbf{B}\right) \boldsymbol{c}. \tag{46}$$

Taking the limit as time goes to infinity, we see that $e^{\mathbf{A}t}$ converges to zero since this is the probability of starting at one of the transient states and arriving at a transient state by time t. Thus,

$$\boldsymbol{\pi}' = \lim_{t\to\infty} \boldsymbol{a}' \left(e^{\mathbf{A}t}\mathbf{A}^{-1}\mathbf{B} - \mathbf{A}^{-1}\mathbf{B}\right) = -\boldsymbol{a}'\mathbf{A}^{-1}\mathbf{B}. \tag{47}$$

where $\boldsymbol{\pi}$ is the vector of probabilities for each possible value of Z.

We can then calculate the mean and variance of Z. So, the mean will be

$$\mu_z = \boldsymbol{\pi}'\boldsymbol{z} \tag{48}$$

where $\boldsymbol{z}$ is the vector of possible values of Z. In the current setting, $\mathbf{z}' = (-1, 0, 1)'$. Thus, the variance can be given as

$$\sigma_z^2 = \boldsymbol{\pi}'\boldsymbol{z_2} - (\boldsymbol{\pi}'\boldsymbol{z})^2 \tag{49}$$

where $\mathbf{z_2}$ consists of the squares of the possible values of Z. In the current setting, $\boldsymbol{z_2}' = (1, 0, 1)'$.

We now have all the necessary information to calculate the asymptotic velocity and effective diffusion, except for the covariance. By taking the derivative of (45) with respect to t and simplifying, we obtain a joint density for Z and τ

$$f(z, t) = \boldsymbol{a}' e^{\mathbf{A}t}\mathbf{B} \tag{50}$$

and use this to calculate the cross moment between Z and τ,

$$\mathbb{E}Z\tau = \boldsymbol{a}' \int_0^\infty t e^{\mathbf{A}t} dt \mathbf{B}\boldsymbol{z} \tag{51}$$

Integrating by parts as we did for the mean of τ, we obtain the covariance

$$\sigma_{Z,\tau} = -\boldsymbol{a}'\mathbf{A}^{-1}\mathbf{B}\boldsymbol{z} - \mu_Z\mu_\tau. \tag{52}$$

With these formulae, we can rapidly assess the dependence of motor run length, velocity and diffusion on particular parameters in the hydrolysis cycle. In previous work, we used standard stochastic simulations of the kinesin hydrolysis cycle to interpret Kinesin-1 and Kinesin-2 run lengths and velocities and identify specific rate constants that determine overall motor characteristics (Muthukrishnan et al. (2009), Shastry and Hancock (2010)). As seen in Figure 2, when k_{attach} is increased from its nominal value (which is partially rate limiting), the run length increases linearly and both the velocity and diffusion increase to a point where attachment is far from rate limiting. In contrast, increasing $k_{\text{hydrolysis}}$ from its nominal value has no effect on the run length, but it has a similar effect on motor velocity and diffusion. While this type of analysis can also be done using the Wang-Peskin-Elston method, we show in the next section that it can be generalized to handle more elaborate examples that include tethered diffusion of the free head (Wang et al. (2003)).

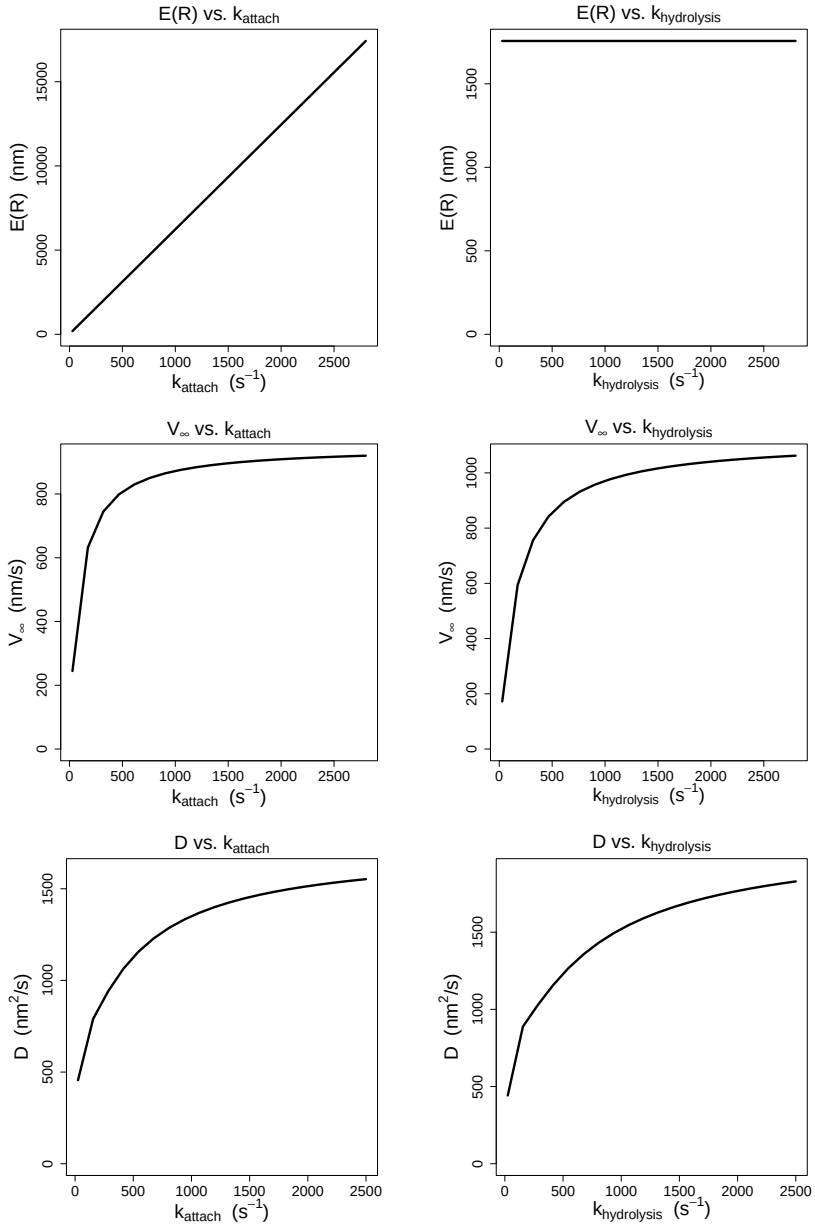
4. The Full Model: Kinetics with a Diffusive Free Head

Processive molecular motors like kinesin walk along a periodically structured microtubule track by transitioning between states where one head is fixed and the other is free and a state in which both heads are bound. Recent experiments in which the neck linker between the two heads is artificially extended have shown that asymptotic velocity, effective diffusion, and run length can all be altered through this modification (Yildiz et al. (2008), Muthukrishnan et al. (2009), Hackney et al. (2003), Shastry and Hancock (2010)).

In general, a motor model should include both the chemical state transitions and the diffusivity of the free head. In Section 3 above, only chemical states were modeled through a discrete state Markov chain. We now present a model that does not depend on the periodic structure *per se* but instead focuses on the cycle consisting of unbinding of one head, the tethered diffusion of that free head, and eventual rebinding; this model was previously explored using stochastic simulation methods (Kutys et al. (2009)).

Assume first that the front head is fixed as in Figure 1. Furthermore, we will shift the problem so that this leading fixed head is located at position

Figure 2: $\mathbb{E}R$, V_∞, and D for the four chemical state model when varying k_{attach} and $k_{\text{hydrolysis}}$ using the values from Table 1.



zero. (The situation when the rear head is fixed is relatively symmetric.) When the rear head detaches, it will have an initial position of $-L$ (see figure 1). The basic model of this tethered diffusion will be a stochastic differential equation that represents the tether of the neck linker and the Brownian fluctuations inherent in the diffusion of a single particle at the nano scale. As before, we are primarily interested in two things: the distribution of the time until the motor becomes bound to a new microtubule binding site and the distribution of the location of that binding site, as well as the asymptotic quantities we can calculate from these distributions.

A simpler SDE model would assume that this binding time is a hitting time to L or $-L$, the binding sites forward and backward, and the probability of a forward or backward steps is merely the probability of arriving to one of these sites or the other. However, considering the true three-dimensional geometry, this may not be reasonable, so we establish a probability of binding when the motor is within a certain radius of the binding site. When a motor head is near to a binding site, an exponential clock will be started (where the rate may depend on the proximity of the head to the binding site). Each binding has an independent clock. If the clock is triggered, then the motor binds to that site. This type of model is similar to the detailed model presented in Atzberger et al and explored through stochastic simulation; however, we are only considering one dimensional dynamics to explore the nature of the neck linker extension(Atzberger and Peskin (2006)).

The position of the free motor head is governed by the following equation.

$$Y(t) = y + \int_0^t a_{M(s)}(Y(s))ds + \sigma B(t) \tag{53}$$

where $M(t)$ is the discrete Markov chain corresponding to chemical events and $B(t)$ is a standard Brownian motion. The drift $a_\cdot(\cdot)$ is determined by the nature of the neck linker tether between the free and bound head and the location of the bound head. The drift can be thought of as the instantaneous mean velocity of the free head. A relatively straightforward example for the drift can be derived from the potential energy corresponding to a Hookean spring representing the tethering of the freely diffusing kinesin head to the bound head. In this case, $a_j(Y(s)) = -k(y - y_0)/\zeta$ where y_0 is an offset, k is the spring constant, and ζ is the drag coefficient. This example along with some alternatives is discussed later in the section.

The binding processes could be written as

$$A_j(t) = N_j\left(\int_0^t g_j(Y(s))ds\right) \tag{54}$$

where the N_j are independent standard Poisson processes (independent of B also). The index j corresponds to a possible binding site. Since the step length for a kinesin motor is 8 nm, then the motor heads are 8 nm apart at the beginning of a cycle. By convention, we will denote the position of our front head as 0 and the rear head as -8. The possible binding sites for the free head will then be -16, -8, 0, and 8 which we could label with $j = 1, 2, 3, 4$ respectively. Note, however, that one site is always blocked by the bound head. For instance if the rear head detaches, then the binding site corresponding to 0 ($j = 3$) will be unavailable for binding. The function $g_j(\cdot)$ is a local binding rate for the site j depending on the position of the free head, $Y(t)$. In general, for a particular position, $Y(t)$, only one of the functions, $g_j(Y(t))$, will be non-zero. One possible selection of the $g_j(\cdot)$ would be a constant near the respective binding sites; another possibility would be a function that increased to a maxima at the binding site. Note that this is equivalent to setting time to absorption in state j as independent random variables, τ_j with the distribution conditioned on Y as

$$P(\tau_j > t) = e^{-\int_0^t g_j(Y(s))ds}. \tag{55}$$

The time to binding is then defined as

$$\tau = \inf\{t : A_j(t) > 0 \ for\ all\ j\} \tag{56}$$

We also define $Y(\tau)$ to be the location of the free head at the end of a cycle.

4.1. An Approximating Markov Chain

We would like to be able to calculate the distribution of τ and Z, the time and distance traveled in this one step. However, this may not always be simple depending on the drift of the diffusion process, $a.(\cdot)$. The most obvious way to calculate these distributions is through Monte Carlo simulations; however, calculating the sensitivity of the output through perturbations of parameters can be both time consuming and lack precision due to sampling variability. One could attempt to calculate the density of $Y(t)$ by solving the underlying Fokker-Planck equation either numerically or analytically or

solving for the hitting time moments directly using deterministic differential equation methods (Karlin and Taylor (1981), Latorre et al. (2007), Elston and Peskin (2000)).

Another relevant numerical method is the Wang, Peskin, and Elston method that relies on the creation of an approximating discrete space, continuous time Markov chain (Wang et al. (2003)). The original context for the WPE method was to consider a purely periodic process and use arguments concerning the stationary distribution of the location of the motor within a cycle. However, the choice of approximating transition rates should work in a more general context. Moreover, the methods presented in this section are similar to previous methods that include the role of cargo in calculations of motor dynamics (Xing et al. (2005),Fricks et al. (2006)). However, those methods continue to rely on a purely periodic structure of the motor dynamics along with an underlying assumption of stationarity of the distance between the cargo and the motor.

Assume, then, that $y_1, ..., y_n$ is an evenly spaced grid on the real numbers with distance between grid points of Δ. We could represent an approximating Markov chain for $dY(t) = a(Y(t))dt + \sigma(Y(t))dB(t)$ using the tridiagonal transition matrix, $\mathbf{L}$, with elements given by

$$\begin{aligned}
\mathbf{L}_{i,i-1} &= \left(\frac{\sigma^2(y_i)}{2} + a^-(y_i)\Delta\right)/\Delta^2 \qquad (57)\\
\mathbf{L}_{i,i+1} &= \left(\frac{\sigma^2(y_i)}{2} + a^+(y_i)\Delta\right)/\Delta^2\\
\mathbf{L}_{i,j} &= 0 \text{ if } |i-j| > 1\\
\mathbf{L}_{i,i} &= -(\mathbf{L}_{i,i-1} + \mathbf{L}_{i,i+1}),
\end{aligned}$$

where $a(y) = a^+(y) - a^-(y)$ is the drift and $\sigma(y)$ is the diffusion coefficient, which should be constant in this case of Brownian diffusion in a potential. We are using the approximating Markov chain rates found in Kushner and Dupuis; as previously mentioned, other choices for approximating rates are also suitable (Kushner and Dupuis (2001)).

We would like to incorporate this model of diffusion into a computational framework that also includes chemical transitions through an approximating Markov chain. Within each chemical state, the diffusion will be determined by a particular drift function, $a_m(y)$. Therefore, we can construct the approximating chain for Y, M by using a block structure. The "outer" structure will describe the discrete chemical reactions, and the "inner" structure will

handle the diffusional model of position of the free head. The block structure will mimic the transitions of the purely diffusive model. In the current case we need to have a common start and stop point to describe a step. A natural choice is the state in which both heads are bound. In the current scenario, state 1_+ will represent only the *initial* state in which both heads are down. State 1_{++} will represent the state to be arrived at after leaving state 1_+ by way of trailing head detachment followed by the trailing head reattaching at two binding sites ahead. State 1_- will represent the state to be arrived at after leaving state 1_+ by way of front-head detachment followed by a binding of this head to the site behind the bound head (i.e. a reverse cycle in Figure 1). As in the kinetic model of the previous section, a null step can also occur as the motor returns to its original conformation, which we denote by 1_*. This scheme will allow us to track the change in the position of the motor between cycles. (We will assume that the position of the motor at the end of the cycle is the location of the front head when both are heads are attached.)

The block form, which is quite similar to the matrices given in Section 3, is as follows.

$$\mathbf{A} = \begin{pmatrix} \mathbf{K}_{1_+,1_+} & \mathbf{K}_{1_+,2_+} & \mathbf{0} & \mathbf{0} & \mathbf{K}_{1_+,4_-} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{2_+,2_+} & \mathbf{K}_{2_+,3_+} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{3_+,2_+} & \mathbf{K}_{3_+,3_+} & \mathbf{K}_{3_+,4_+} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_{4_+,3_+} & \mathbf{K}_{4_+,4_+} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{K}_{4_-,4_-} & \mathbf{K}_{4_-,3_-} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{K}_{-3,-4} & \mathbf{K}_{3_-,3_-} & \mathbf{K}_{3_-,2_-} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{K}_{2_-,3_-} & \mathbf{K}_{2_-,2_-} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{58}$$

and

$$\mathbf{B} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{2_+,1_*} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{K}_{4_+,1_{++}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{4_-,1_*} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_{2_-,1_-} \end{pmatrix}. \tag{59}$$

A block in matrix $\mathbf{A}$ generally has one of three forms: a zero matrix, a diagonal matrix, or a tridiagonal matrix. Each block is $n \times n$, where n is the

number of points in the spatial grid. We suggest that the grid run from -24 nm to 16 nm. Note that the grid must contain the locations of the binding sites, namely, -16 nm, -8 nm, 0 nm, and 8 nm.

Note that these newly defined matrices can be arranged as before into a $\mathbf{Q}$ matrix. We will denote each sub-matrix of this matrix as $\mathbf{Q}_{i,j}$. Each of these sub-matrices will be described below; the right way to think about such a sub-matrix is representing the state transitions between spatial grid points. For example, the matrix $\mathbf{K}_{2_+,2_+}$ is a tridiagonal matrix with transition rates corresponding to (57). Note, however, that the diagonal will be different here—it will be constrained so that the rows of $\mathbf{Q}_{2,\bullet}$ sum to 0. Since the rear head detached, the potential should be centered behind the forward head which is bound, we will assume that this bias is -4 nm. (We are following the modeling assumptions presented in Kutys et al. (2009); other choices could be made.) Hence, $a(y) = -k_2(y + 4)$, where k_2 is the spring constant. This can be represented in the following way:

$\mathbf{Q}_{2,2} = \mathbf{K}_{2_+,2_+} =$

$$\begin{array}{c} \\ y_1 \\ y_2 \\ \dots \\ \dots \\ y_n \end{array} \begin{array}{c} \begin{array}{cccccc} y_1 & y_2 & \dots & \dots & \dots & y_n \end{array} \\ \begin{pmatrix} -\sum_{n+1} & \mathbf{L}_{1,2} & 0 & 0 & \dots & 0 \\ \mathbf{L}_{2,1} & -\sum_{n+2} & \mathbf{L}_{2,3} & 0 & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \mathbf{L}_{n-2,n-1} & -\sum_{n+(n-1)} & \mathbf{L}_{n-1,n} \\ 0 & \dots & \dots & 0 & \mathbf{L}_{n,n-1} & -\sum_{2n} \end{pmatrix} \end{array}$$

where $\sum_i$ is the sum of all non-diagonal elements of the ith row of matrix $\mathbf{Q}$. The $\mathbf{L}_{\bullet,\bullet}$ entries are as defined above using the local approximation of the diffusion process.

The entries of matrix $\mathbf{K}_{1_+,2_+}$ are 0 except for the column corresponding to spatial location -8 nm. This column contains the rate at which the motor's rear head becomes detached, k_{detach}. Similarly, $\mathbf{K}_{1_+,4_-}$ is zero except for the column corresponding to spatial location 0, and the nonzero column contains k'_{attach}, the rate at which the front head becomes detached. Since the remaining off-diagonal elements of $\mathbf{Q}_{1,\bullet}$ are zero matrices, $\mathbf{K}_{1_+,1_+}$ is diagonal: $\mathbf{K}_{1_+,1_+} = \text{diag}(-(k_{\text{detach}} + k'_{\text{attach}}))$.

$\mathbf{K}_{2_+,3_+}$ is a diagonal matrix with k_{on} on its diagonal, k_{on} being the rate at which the motor transitions from chemical state 2_+ to chemical state 3_+ through binding of ATP to the front head. The matrix is diagonal because a chemical change leaves the free head's position unchanged; however the center of the potential (i.e. drift function) is shifted as described below.

The matrices of rows 3 and 4 are similar to those of row 2. The matrices $\mathbf{K}_{3_+,2_+}$ and $\mathbf{K}_{4_+,3_+}$ are diagonal matrices with transition rates k'_{on} and $k_{\text{hydrolysis}}$, respectively. $\mathbf{K}_{3_+,3_+}$ is similar to $\mathbf{K}_{2_+,2_+}$ but with a different drift function. Here the bias should be forward so that $a(y) = -k_3(y-4)$, i.e., the center of the potential is located midway between the bound head and the next binding site. This transition corresponds to docking of the neck linker, which is thought to be the principal conformational change in the kinesin mechanochemical cycle (Rice et al. (1999)). State 4 has the same drift function, perhaps with a different spring constant: $a(y) = -k_4(y-4)$. $\mathbf{K}_{4_+,4_+}$ and $\mathbf{K}_{3_+,3_+}$ have the same structure as $\mathbf{K}_{2_+,2_+}$.

For the forward cycle, it is left to describe $\mathbf{K}_{2_+,1_*}$ and $\mathbf{K}_{4_+,1_+}$, which correspond to absorption in state 1_{++} resulting in the leading head being 8 nm ahead of where it was in the last cycle. When the front head binds it may "jump" from a nearby location to the binding site when the respective binding process fires. From state 2_+, for example, the free head can bind at location -8 nm or location 8 nm corresponding to states 1_* and 1_{++} respectively. If location -8 nm corresponds to process N_2 and location 8 nm corresponds to process N_4, matrix $\mathbf{K}_{2_+,1_*}$ and $\mathbf{K}_{4_+,1_+}$ zero with the exception of the columns corresponding to those binding locations. The columns consist of $g_j(y_i)$ for $j = 2$ or $j = 4$ respectively with $i = 1, \ldots, n$ where $g_j(\cdot)$ is defined above and y_i is the grid point corresponding to row i. The structure of $\mathbf{K}_{4_+,1_{++}}$ is similar. Note that binding at 8 nm from state 2_+ or binding at -8 nm from state 4_+ is unlikely and the rates are therefore not included. In addition, we make a modeling assumption that the binding of the free head cannot occur in state 3_+; there is nothing in our framework preventing such a transition.

The submatrices for the back cycles are similar, but the geometry is slightly different; since the front head has detached, the bound head is located at -8 nm. The matrices $\mathbf{K}_{4_-,3_-}$, $\mathbf{K}_{3_-,4_-}$, $\mathbf{K}_{3_-,2_-}$, and $\mathbf{K}_{2_-,3_-}$ are identical to their counterparts in the forward cycle. Matrices $\mathbf{K}_{3_-,3_-}$ and $\mathbf{K}_{4_-,4_-}$ have the same structure as $\mathbf{K}_{3_+,3_+}$, but the drift is now centered at -4 nm. $\mathbf{K}_{2_-,2_-}$ is the same as $\mathbf{K}_{2,2}$, but its drift is now centered at -12 nm. Finally, matrices $\mathbf{K}_{4_-,1_*}$ and $\mathbf{K}_{2_-,1_-}$ are identical in form to $\mathbf{K}_{4_+,1_{++}}$ and $\mathbf{K}_{2_+,1_*}$. Now that we have formulated the $\mathbf{A}$ and $\mathbf{B}$ matrices, we can use the formulae given in the previous section to compute various probabilities and moments.

Because it does not account for dissociation, the form of $\mathbf{B}$ given above is suitable for computing the asymptotic quantities V_∞ and D. To compute $\mathbb{E}R$, however, we must provide an additional absorbing state, call it state $\emptyset$ that represents dissociation. This requires only that we add a column of

submatrices to $\mathbf{B}$ to arrive at

$$\mathbf{B} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{2_+,1_*} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{K}_{4_+,1_{++}} & \mathbf{0} & \mathbf{0} & \mathbf{K}_{4_+,\emptyset} \\ \mathbf{0} & \mathbf{K}_{4_-,1_*} & \mathbf{0} & \mathbf{K}_{4_-,\emptyset} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_{2_-,1_-} & \mathbf{0} \end{pmatrix}. \tag{60}$$

where $\mathbf{K}_{4,6} = \mathbf{K}_{-4,6} = \mathrm{diag}(k_{\mathrm{unbind}})$.

4.2. Modeling Neck Linker Extension

In a previous paper, we considered three candidate models for the neck linker when one head is diffusing: Hookean, Worm Like Chain (WLC), and Finitely Elastic Nonlinear Extensible (FENE) (Kutys et al. (2009)). Exploring motor behavior using different representations of the neck linker is important because while it is known that the neck linker plays a key role in kinesin stepping, the underlying details of tethered diffusion are poorly understood. Any model for the neck linker enters into the above mentioned framework through the drift function, $a(y)$, and so our procedure offers a numerical technique for investigating the effect of varying one or more parameters on motor behavior—asymptotic velocity, effective diffusion, expected run length.

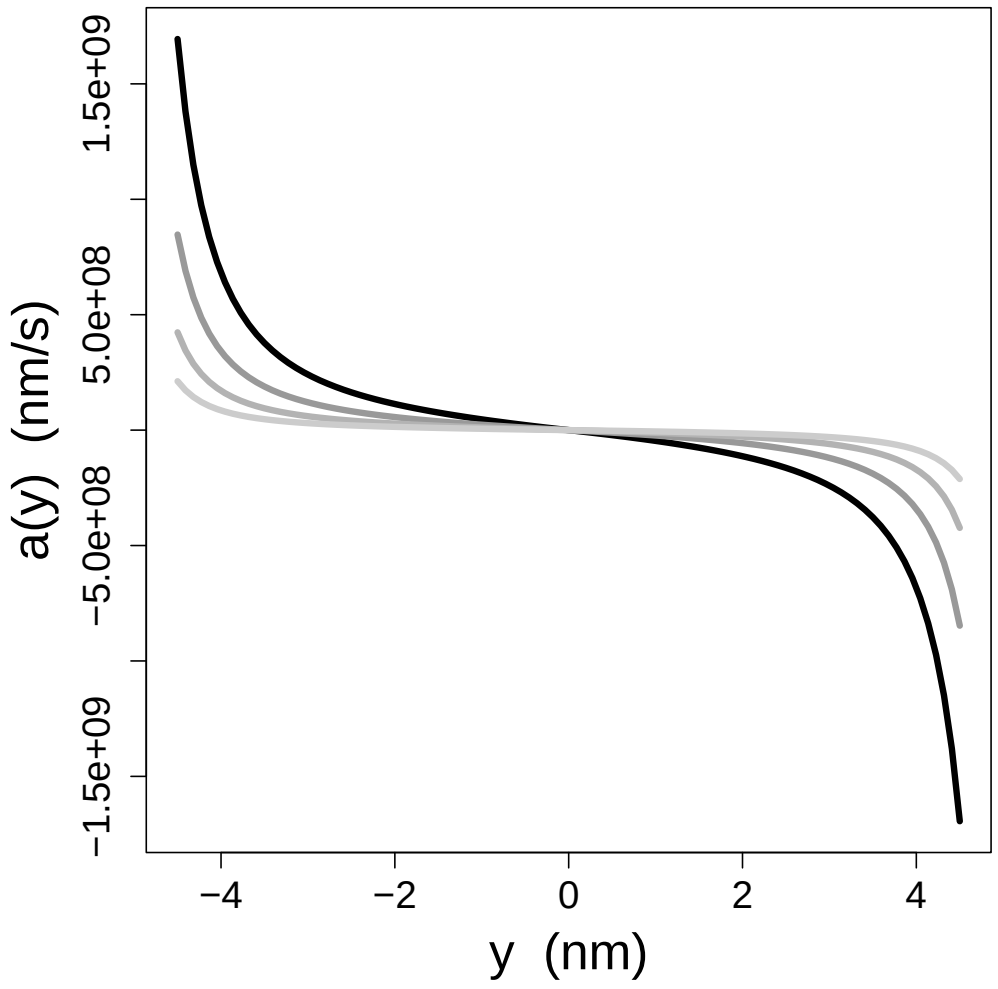
Perhaps the simplest of the three neck linker models is the Hookean model, so named because it considers the neck linker to be a Hookean spring. This implies the drift function $a(y) = -k(y - y_0)/\zeta$, where k is the spring constant (we use 1 pN/nm as the default value), y_0 is the center of the potential, and $\zeta = 5.66 \times 10^{-8} pN\cdot$ sec/nm is the frictional drag coefficient. This model was used in the preceding section to describe the construction of $\mathbf{Q}$.

In the WLC scenario, the neck linker is modeled as becoming increasingly stiff with extension up to a fixed limit known as the contour length, L_c. The WLC drift is given by

$$a(y) = -\operatorname{sign}(y - y_0)\frac{k_B T}{\zeta L_p}\left(\frac{1}{4}\left(1 - \frac{|y - y_0|}{L_c}\right)^{-2} - \frac{1}{4} + \frac{|y - y_0|}{L_c}\right) \text{ if } |y - y_0| < L_c, \tag{61}$$

where $k_B T$ is the Boltzmann constant times absolute temperature and L_p is the persistence length for a chain of amino acids. A plot of this drift (for $L_p = 0.5$ nm, $y_0 = 0$ nm, and $L_c = 5.3$ nm) is shown in Figure 3.

Figure 3: The WLC drift function. Here the drift function is centered at 0, $L_c = 5.3$ nm, and four values of L_p: 0.5 nm(black), 1 nm, 2 nm, and 4 nm (lightest gray).



Parameter	Default Value
k (Hookean)	1 pN/nm
k (FENE)	0.01 pN/nm
ζ	5.66×10^{-8} pN nm s^{-1}
L_p	0.5 nm
L_c	5.3 nm
σ^2	1.46×10^8 nm^2 s^{-1}
binding radius	1 nm
k_{attach} (Hookean)	8,000 s^{-1}
k_{attach} (WLC)	180,000 s^{-1}
k_{attach} (FENE)	2,800 s^{-1}
k_{detach}	250 s^{-1}
k'_{detach}	0.1 s^{-1}
k_{on}	2,000 s^{-1}
k'_{on}	200 s^{-1}
$k_{\text{hydrolysis}}$	280 s^{-1}
$k'_{\text{hydrolysis}}$	3.5 s^{-1}
k_{unbind}	1.7 s^{-1}

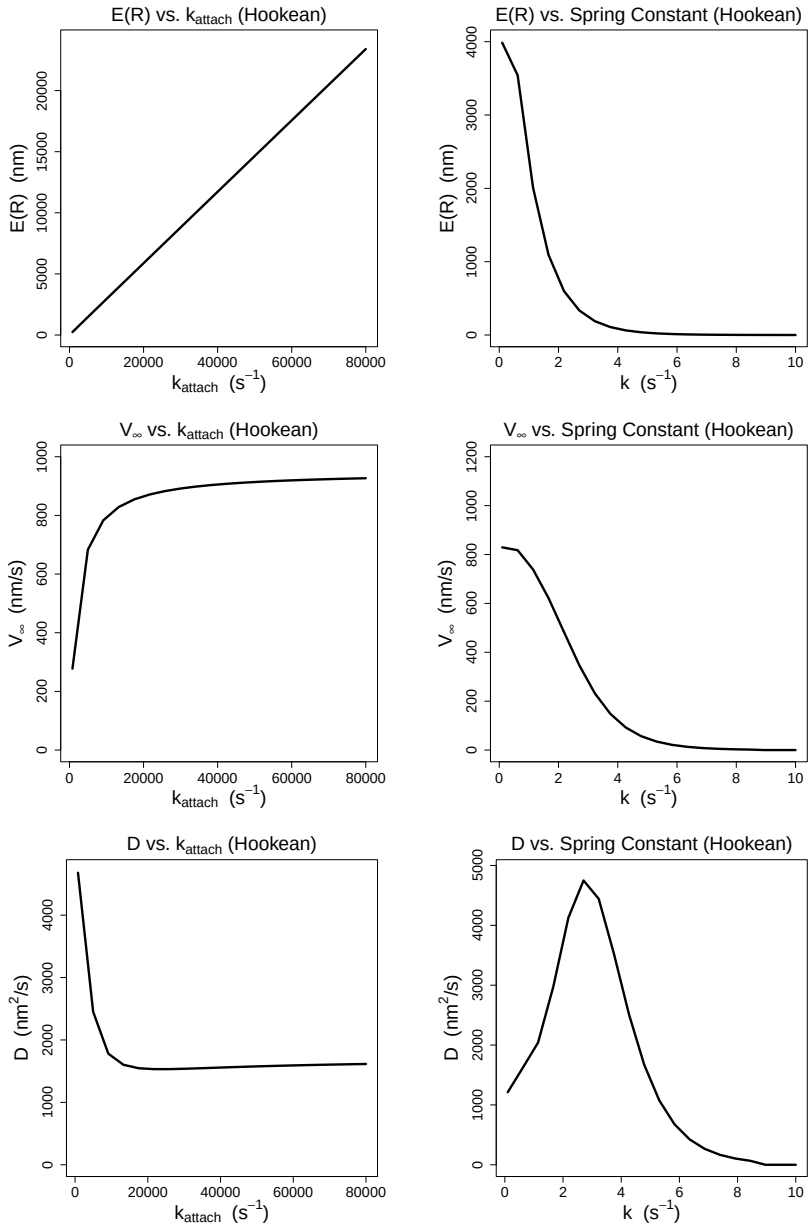
Table 1: Default parameter values. Note that k_{on} depends on the concentration of ATP. Specifically, we are assuming $k_{\text{on}} = k_{\text{on}}^{ATP}[ATP]$ where $k_{\text{on}}^{ATP} = 2\mu M^{-1}$ s^{-1} and $[ATP] = 1000\mu M$. The notation k' denotes the reverse kinetic rate corresponding to k.

The FENE model posits a neck linker that allows the free head to diffuse quite freely but only in $(y_0 - L_c, y_0 + L_c)$. The corresponding drift function is $a(y) = -k(y - y_0)/\zeta$ if $|y - y_0| < L_c$, where k is a small spring constant (e.g., 0.01 pN/nm). Conceptually, this drift increases dramatically as the displacement from y_0 increases to L_c, where it abruptly asymptotes. In practice, the transition rates for the approximating Markov chain are set to zero, thus preventing movement outside the set boundaries.

Table 1 gives the default parameter values mentioned above along with our default values for the chemical rates. Note that the default attachment rates are different for the three models. Because this is a local attachment rate instead of a typical kinetic rate, little is known about the value. The values given here were selected to guarantee a experimentally realistic velocity and an effective attachment rate of approximately 600 s^{-1}. For more detail, see Kutys et al. (2009).

As a first test of the method, we use the simplest model, the Hookean

Figure 4: $\mathbb{E}R$, V_∞, and D for the Hookean model plotted as a function of the attachment rate of the free head while within the binding radius, k_{attach}, and κ, the spring constant.

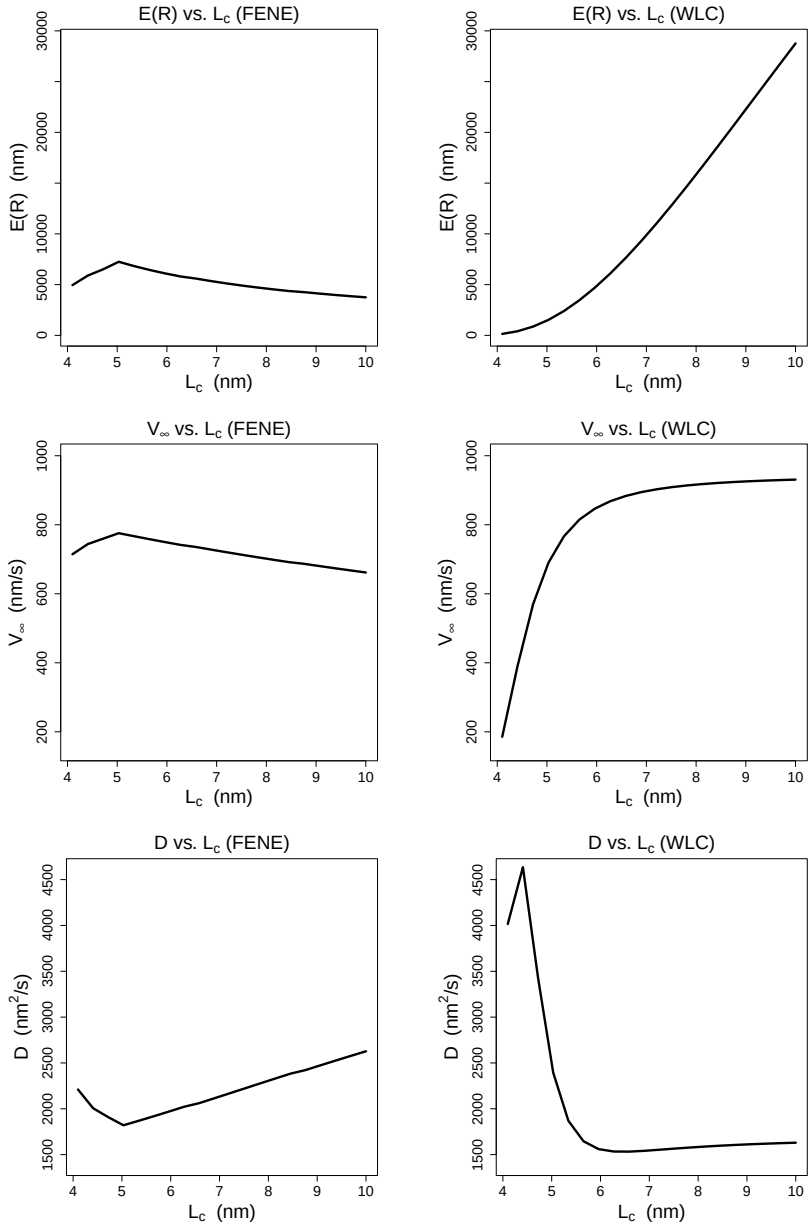


spring, and examine the dependency of relevant motor metrics on two variables, the spring constant and the attachment rate for the free head while within the binding radius. In Figure 4, we see that the expected run length increases linearly with the attachment rate, k_{attach}. This is not surprising given that detachment can occur only in this state (see Figure 1); changing this rate will shorten the percentage of time spent in this detachable state, thus increasing run length. However, changing the spring constant has a much stronger effect. As the spring is stiffened, the free head will require more time to "search" for the binding site given its mechanical constraint. This makes detachment during each cycle more likely, causing the expected run length to go down. Increasing the hydrolysis rate has a stronger effect on velocity and effective diffusion. The velocity is increased because the motor spends less time in state 3 before ATP is hydrolyzed, a necessary step to complete the cycle. The increased diffusion can be explained by the Poisson nature of this purely chemical step.

One question inspired by the neck linker extension experiments is how qualitatively different models of the neck linker respond to extension. In Figure 5, we see the quantities of interest for the FENE and WLC models as the contour length (total extension length) changes. We see that as L_c is enlarged the FENE model shows a mild decrease in run length after an initial brief increase. This is strongly contrasted by the rapid increase in run length with increasing L_c for the WLC model. The experimental evidence is contradictory on the run length change as a function of L_c with Hackney and Mutukrishnan et al separately showing a decrease in run length while Yildiz found no change (Yildiz et al. (2008), Muthukrishnan et al. (2009), Hackney et al. (2003)). The relative insensitivity of the FENE model could be the reason for the difficulty in detection, but the WLC seems to contradict these experimental results.

The velocity in the FENE model also decreases with increasing L_c, which is again confirmed by the experimental evidence. The WLC shows a rather rapid increase in velocity with increasing L_c as well. As the FENE model's region of exploration increases, the time required to search over the relevant range and find the binding site increases, slowing the overall cycle time. In contrast, for the WLC model, the tight constraint means that the motor head can reach the binding site only rarely, and extending the neck linker allows the free head to reach the binding site more readily. The result is that extending L_c in the FENE model lengthens the search time and increases the effective diffusion, while extending L_c in the WLC model relieves the constraints of the

Figure 5: $\mathbb{E}R$, V_∞, and D versus contour length, L_c, for the FENE and WLC models.



search, shortening the search time and leading to a decrease in the effective diffusion.

In our model, we have both diffusive and biochemical dynamics. In some cases these may be simple to combine; for instance, where we assume that the hydrolysis step does not depend on the diffusion of the free head. However, a more complicated situation arises when considering the binding of the free head to the microtubule. A classical approach is to simply consider the hitting time of the free head to the binding site location, but this does not consider such complications as the proper rotational orientation of the head with respect to the microtubule. While there are experimental data for ADP release (which is catalyzed by the motor head binding to the microtubule), there is little accessible data for the kinetics of the motor head binding to the microtubule or more generally binding in proximity (Hackney (2002), Zarnitsyna et al. (2007), Guydosh and Block (2009)). Because of this, the local binding rates and the binding radius were somewhat arbitrarily chosen. However, within this framework, we can easily explore the dependency of the system on these rates. As seen in Figure 6, the FENE model shows a steady increase in run length when the binding radius and the binding rate are increased. The reason is that increasing either of these quantities will decrease the time spent in the "search" state. The WLC model shows a more non-linear sensitivity to the radius. This is not too surprising given the increasing force felt by the free head as it approaches full extension; this plot shows that there is a rapid increase in run length as this radius increases.

For velocity, both models show a fairly strong increase in velocity as the radius is increased. This is especially true of the WLC model where the divide between the range where most of the head's time is spent is particularly sharp. The velocity of the WLC model is somewhat insensitive to the binding rate, however. One explanation is that the difficulty for the WLC to bind is in arriving at the binding site–not in the binding rate itself.

In Figure 7, we see the run length, velocity, and effective diffusion plotted against the contour length for various level of persistence length. As the persistence length is decreased, the tether loosens to the finite extension and thus qualitatively approaches the FENE model. The curves begin to approach those of the FENE model of Figure 5.

Figure 6: $\mathbb{E}R$, V_∞, and D versus binding radius for the FENE and WLC models. Each panel shows curves for four values of k_{attach}: default/4 (black), default/2, default, $2 \cdot$ default, and $4 \cdot$ default (lightest gray). Default k_{attach} is $2800s^{-1}$ for the FENE model and $180,000$ for the WLC model.

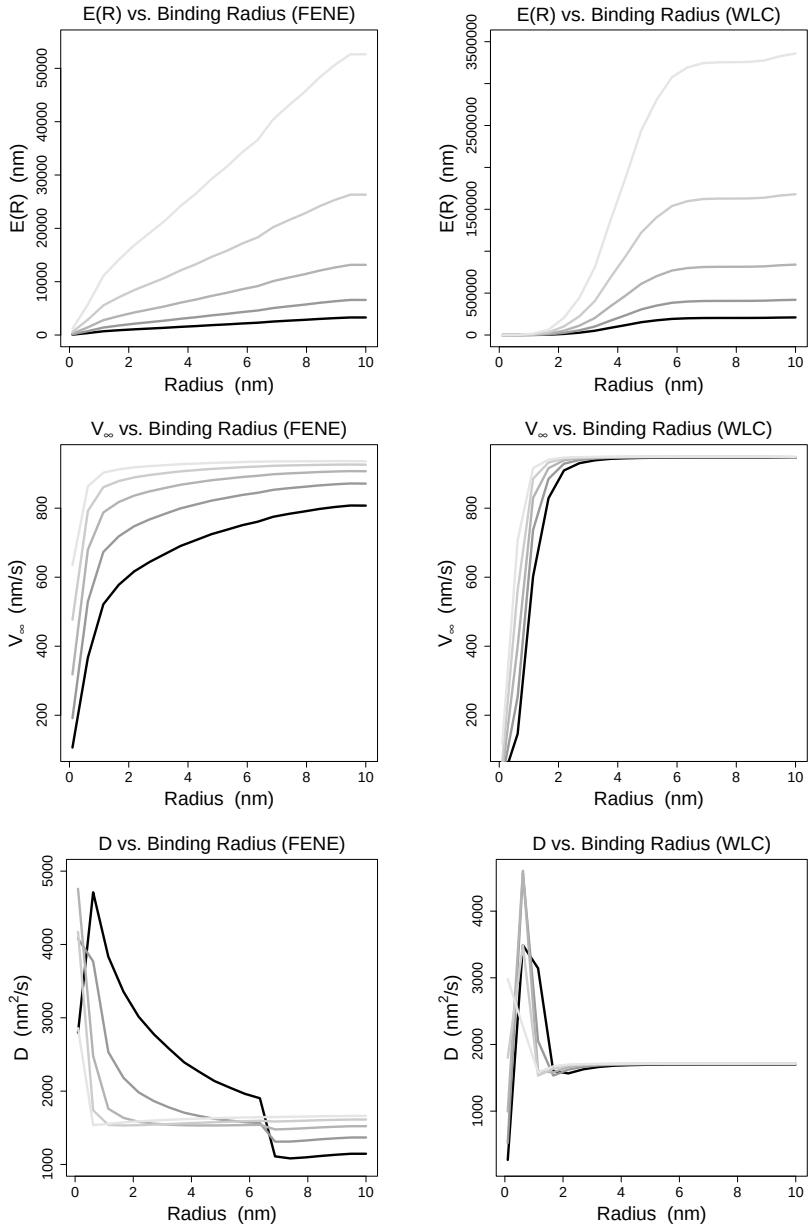
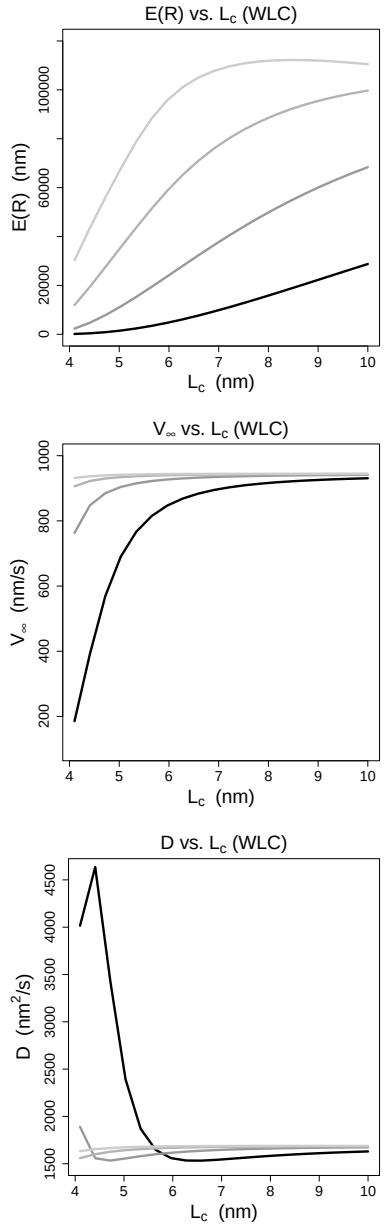


Figure 7: $\mathbb{E}R$, V_∞, and D versus contour length for the WLC model. Each panel shows curves for four values of L_p: 0.5nm(black), 1nm, 2nm, and 4nm (lightest gray).



5. Conclusion

In this work, we have presented a methodology for numerical exploration of neck linker extensions for kinesin molecular motors. By decomposing the movement of the motor into steps and using approximating Markov chains, we are able to present fairly straightforward matrix calculations that can quickly return relevant experimentally measured values for a given model for comparison to *in vitro* experiments. Given the large parameter space for these models, these methods allow for rapid exploration giving better insights into biologically relevant aspects of a given model. This extends the work by Wang et al by giving a matrix formulation which includes explicit information about the individual steps of the motor through the distribution of the step time and the direction of the step (Wang et al. (2003)). In addition, the application of the method presented in the paper allows the explicit modeling of both heads of the kinesin linking these local dynamics to the movement of the motor along the microtubule.

6. Acknowledgments

We gratefully acknowledge the NSF who supported the present work though the Joint DMS/NIGMS Initiative to Support Research in the Area of Mathematical Biology (DMS-0714939) and through a visiting position at Statistical and Applied Mathematical Sciences Institute for John Fricks.

7. Appendix

In this section, we derive the formula for the effective diffusion of the stepping model when there is dependence between the step direction and the stepping time.

Define

$$S(t) = \sum_{i=0}^{\lfloor t \rfloor} Z_i \quad T(t) = \sum_{i=0}^{\lfloor t \rfloor} \tau_i \tag{62}$$

where $\lfloor t \rfloor$ is the nearest smallest integer to t.

Now, the central limit theorem for random vectors will yield the following functional central limit theorem:

$$\begin{pmatrix} X_n(t) \\ Y_n(t) \end{pmatrix} = n^{-1/2} \begin{pmatrix} S(nt) - \mu_Z nt \\ T(nt) - \mu_\tau nt \end{pmatrix} \Rightarrow \begin{pmatrix} B_x(t) \\ B_y(t) \end{pmatrix} \tag{63}$$

where the two dimensional Brownian motion on the right has covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_Z^2 & \sigma_{Z,\tau} \\ \sigma_{Z,\tau} & \sigma_\tau^2 \end{pmatrix}. \tag{64}$$

Now, if we define

$$U_n(t) = n^{-1/2}\left(S(T^{-1}(nt)) - \frac{\mu_Z}{\mu_\tau}nt\right) \tag{65}$$

and we apply Theorem 13.7.3 from Whitt (2002), we obtain

$$U_n(t) \Rightarrow B_x\left(\frac{t}{\mu_\tau}\right) - \frac{\mu_Z}{\mu_\tau}B_y\left(\frac{t}{\mu_\tau}\right). \tag{66}$$

The RHS of (66) could be rewritten as

$$\begin{pmatrix} 1 & -\frac{\mu_Z}{\mu_\tau} \end{pmatrix} \begin{pmatrix} B_x\left(\frac{t}{\mu_\tau}\right) \\ B_y\left(\frac{t}{\mu_\tau}\right) \end{pmatrix}. \tag{67}$$

Since the limit process is a linear combination of Brownian motions, it is also a Brownian motion with

$$2D = \begin{pmatrix} \frac{1}{\sqrt{\mu_\tau}} & -\frac{\mu_Z}{\mu_\tau^{3/2}} \end{pmatrix} \Sigma \begin{pmatrix} \frac{1}{\sqrt{\mu_\tau}} & -\frac{\mu_Z}{\mu_\tau^{3/2}} \end{pmatrix}' \tag{68}$$

Note that we have also applied the standard change of time for Brownian motion–$B(ct)$ is equivalent in law to $c^{1/2}W(t)$. If we expand this expression, we get that the limiting process in 66 is a Brownian motion with

$$2D = \frac{\sigma_Z^2}{\mu_\tau} + \frac{\mu_z^2\sigma_\tau^2}{\mu_\tau^3} - 2\frac{\mu_Z\sigma_{Z,\tau}}{\mu_\tau^2}. \tag{69}$$

References

Asmussen, S., 2003. Applied probability and queues. Springer Verlag.

Atzberger, P., Peskin, C., 2006. A brownian dynamics model of kinesin in three dimensions incorporating the force-extension profile of the coiled-coil cargo tether. Bulletin of mathematical biology 68, 131–160.

Billingsley, P., 2008. Probability and measure. Wiley.

Block, S., 2007. Kinesin motor mechanics: binding, stepping, tracking, gating, and limping. Biophysical Journal 92, 2986–2995.

Block, S.M., Goldstein, L.S., Schnapp, B.J., 1990. Bead movement by single kinesin molecules studied with optical tweezers. Nature 348, 348–52.

Colquhoun, D., Hawkes, A., 1982. On the stochastic properties of bursts of single ion channel openings and of clusters of bursts. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 300, 1–59.

Cox, D., 1962. Renewal Theory. Chapman & Hall.

Cross, R., 2004. The kinetic mechanism of kinesin. Trends in Biochemical Sciences 29, 301–309.

Das, R., Kolomeisky, A., 2009. Dynamic properties of molecular motors in the divided-pathway model. Physical Chemistry Chemical Physics 11, 4815–4820.

Elston, T., Peskin, C., 2000. The role of protein flexibility in molecular motor function: coupled diffusion in a tilted periodic potential. SIAM Journal on Applied Mathematics 60.

Fredkin, D., Rice, J., 1986. On aggregated markov processes. Journal of Applied Probability , 208–214.

Fricks, J., Wang, H., Elston, T., 2006. A numerical algorithm for investigating the role of the motor–cargo linkage in molecular motor-driven transport. Journal of Theoretical Biology 239, 33–48.

Guydosh, N., Block, S., 2009. Direct observation of the binding state of the kinesin head to the microtubule. Nature 461, 125–128.

Hackney, D., 2002. Pathway of ADP-stimulated ADP release and dissociation of tethered kinesin from microtubules: implications for the extent of processivity. Biochemistry 41, 4437–4446.

Hackney, D., Stock, M., Moore, J., Patterson, R., 2003. Modulation of kinesin half-site adp release and kinetic processivity by a spacer between the head groups. Biochemistry 42, 12011–12018.

Hancock, W., Howard, J., 2003. Molecular Motors. Wiley-VCH, Weinheim, Germany. chapter Kinesin: processivity and chemomechanical coupling. pp. 243–269.

Hariharan, V., Hancock, W., 2009. Insights into the mechanical properties of the kinesin neck linker domain from sequence analysis and molecular dynamics simulations. Cellular and Molecular Bioengineering 2, 177–189.

Hirokawa, N., Pfister, K.K., Yorifuji, H., Wagner, M.C., Brady, S.T., Bloom, G.S., 1989. Submolecular domains of bovine brain kinesin identified by electron microscopy and monoclonal antibody decoration. Cell 56, 867–78.

Julicher, F., Ajdari, A., Prost, J., 1997. Modeling molecular motors. Reviews of Modern Physics 69, 1269–1282.

Karlin, S., Taylor, H., 1975. A first course in stochastic processes. Academic Press New York.

Karlin, S., Taylor, H., 1981. A second course in stochastic processes. Academic press.

Kolomeisky, A., Fisher, M., 2007. Molecular motors: a theorist's perspective. Annual Review of Physical Chemistry 58, 675–695.

Kushner, H., Dupuis, P., 2001. Numerical methods for stochastic control problems in continuous time. Springer Verlag.

Kutys, M., Fricks, J., Hancock, W., 2009. Monte carlo analysis of neck linker extension in kinesin molecular motors. Submitted.

Latorre, J., Kramer, P., Pavliotis, G., 2007. Effective transport properties for flashing ratchets using homogenization theory. Proceedings in Applied Mathematics and Mechanics 7.

Lindner, B., Kostur, M., Schimansky-Geier, L., 2001. Optimal diffusive transport in a tilted periodic potential. Fluctuation and Noise Letters 1, R25–R39.

Mallik, R., Carter, B.C., Lex, S.A., King, S.J., Gross, S.P., 2004. Cytoplasmic dynein functions as a gear in response to load. Nature 427, 649–52.

Miki, H., Setou, M., Kaneshiro, K., Hirokawa, N., 2001. All kinesin superfamily protein, KIF, genes in mouse and human. Proc Natl Acad Sci U S A 98, 7004–11.

Mogilner, A., Wang, H., Elston, T., Oster, G., 2002. Molecular motors: theory and experiment. Computational Cell Biology. C. Fall, E. Marland, J. Wagner, and J. Tyson, editors. Springer-Verlag, New York .

Muthukrishnan, G., Zhang, Y., Shastry, S., Hancock, W., 2009. The processivity of kinesin-2 motors suggests diminished front-head gating. Current Biology 19, 442–447.

Neuts, M., 1994. Matrix-geometric solutions in stochastic models: an algorithmic approach. Dover Pubns.

Rice, S., Lin, A., Safer, D., Hart, C., Naber, N., Carragher, B., Cain, S., Pechatnikova, E., Wilson-Kubalek, E., Whittaker, M., et al., 1999. A structural change in the kinesin motor protein that drives motility. Nature 402, 778–784.

Rock, R.S., Ramamurthy, B., Dunn, A.R., Beccafico, S., Rami, B.R., Morris, C., Spink, B.J., Franzini-Armstrong, C., Spudich, J.A., Sweeney, H.L., 2005. A flexible domain is essential for the large step size and processivity of myosin VI. Mol Cell 17, 603–9.

Schief, W., Howard, J., 2001. Conformational changes during kinesin motility. Current Opinion in Cell Biology 13, 19–28.

Sellers, J.R., Veigel, C., 2006. Walking with myosin V. Curr Opin Cell Biol 18, 68–73.

Shastry, S., Hancock, W., 2010. Neck Linker Length Determines the Degree of Processivity in Kinesin-1 and Kinesin-2 Motors. Current Biology .

Vale, R., Milligan, R., 2000. The way things move: looking under the hood of molecular motor proteins. Science 288, 88–95.

Vale, R.D., Funatsu, T., Pierce, D.W., Romberg, L., Harada, Y., Yanagida, T., 1996. Direct observation of single kinesin molecules moving along microtubules. Nature 380, 451–3.

Wang, H., Peskin, C., Elston, T., 2003. A Robust Numerical Algorithm for Studying Biomolecular Transport Processes. Journal of Theoretical Biology 221, 491–511.

Wang, H., Qian, H., 2007. On detailed balance and reversibility of semimarkov processes and single-molecule enzyme kinetics. J. Math. Phys. 48, 013303.

Whitt, W., 2002. Stochastic-process limits: an introduction to stochastic-process limits and their application to queues. Springer Verlag.

Xing, J., Wang, H., Oster, G., 2005. From continuum fokker-planck models to discrete kinetic models. Biophysical Journal 89, 1551–1563.

Yang, J.T., Laymon, R.A., Goldstein, L.S., 1989. A three-domain structure of kinesin heavy chain revealed by DNA sequence and microtubule binding analyses. Cell 56, 879–89.

Yildiz, A., Tomishige, M., Gennerich, A., Vale, R., 2008. Intramolecular strain coordinates kinesin stepping behavior along microtubules. Cell 134, 1030–1041.

Zarnitsyna, V., Huang, J., Zhang, F., Chien, Y., Leckband, D., Zhu, C., 2007. Memory in receptor–ligand-mediated cell adhesion. Proceedings of the National Academy of Sciences 104, 18037.