

## A STOCHASTIC COMPARTMENTAL MODEL FOR FAST AXONAL TRANSPORT\*

LEA POPOVIC<sup>†</sup>, SCOTT A. MCKINLEY<sup>‡</sup>, AND MICHAEL C. REED<sup>§</sup>

**Abstract.** In this paper we develop a probabilistic microscale compartmental model and use it to study macroscale properties of axonal transport, the process by which intracellular cargo is moved in the axons of neurons. By directly modeling the smallest scale interactions, we can use recent microscopic experimental observations to infer all the parameters of the model. Then, using techniques from probability theory, we compute asymptotic limits of the stochastic behavior of individual motor-cargo complexes, while also characterizing both equilibrium and nonequilibrium ensemble behavior. We use these results in order to investigate three important biological questions: (1) How homogeneous are axons at stochastic equilibrium? (2) How quickly can axons return to stochastic equilibrium after large local perturbations? (3) How is our understanding of delivery time to a depleted target region changed by taking the whole cell point of view?

**Key words.** axonal transport, molecular motor, homogeneity, stochastic process, compartmental model

**AMS subject classifications.** 60, 92

**DOI.** 10.1137/090775385

**1. Introduction.** In all cells, one finds that proteins, membrane-bound organelles, and other structures (e.g., chromosomes) are transported from place to place at speeds much higher than diffusion. Though these transport processes are fundamental to cell function, many of the underlying mechanisms, organizational principles, and regulatory features remain unknown. Axonal transport is one of the best studied systems because the transport is basically one-dimensional since axons are long and narrow. There are two speeds of axonal transport. Fast transport goes at speeds of roughly 0.2 to 0.5 meters/day [26], [32], while slow transport goes at approximately 1 millimeter/day, the rate of axon growth and regeneration [6], [26]. The biology and principles of slow transport are not yet clear [6], but the basic mechanisms of fast axonal transport were discovered in the 1980s [2], [3], [28], [42]. The model in this paper refers to fast axonal transport, which we will henceforth call axonal transport.

The axonal transport apparatus consists of vesicles which form reversible chemical bonds with motor proteins that bind reversibly to microtubules which run parallel to the long dimension of the axon [1], [45]. When the vesicle-motor protein complex is assembled on the microtubule, the complex steps stochastically with step size approximately 8 nanometers for kinesin and dynein and 10 nanometers for myosin [7], [11], [17], [39]. The vesicles enter from the cell body on microtubules and then detach and reattach to the transport mechanism at random times.

In this paper we propose a spatial Markov-chain compartmental model based on

---

\*Received by the editors October 29, 2009; accepted for publication (in revised form) April 15, 2011; published electronically August 25, 2011. This work was supported by NSF grants DMS-061670 and EF-1038593 and by an NSERC Discovery Grant.

<http://www.siam.org/journals/siap/71-4/77538.html>

<sup>†</sup>Department of Mathematics and Statistics, Concordia University, Montreal QC H3G 1M8, Canada (lpopovic@mathstat.concordia.ca).

<sup>‡</sup>Department of Mathematics, University of Florida, 358 Little Hall Box 118105, Gainesville, FL 32605 (scott.mckinley@ufl.edu).

<sup>§</sup>Department of Mathematics, Duke University, Box 90320, Durham, NC 27701 (reed@math.duke.edu).

these dynamics. We will assume independence of the interactions and exponential wait times between events. While we address the validity of these assumptions in the discussion section, we consider this a useful “first-order” approximation that permits study of the dynamics from both the perspective of individual vesicles as well as that of the full spatial system. Such a model unifies all earlier deterministic and stochastic modeling efforts and can accommodate both qualitative and quantitative experimental data observed on multiple scales.

In much experimental work in the 1970s and 1980s, radio-labeled amino acids were put into the cell bodies continuously or for a few hours. The amino acids were incorporated into proteins that were packaged into vesicles and put on the transport system so that at later times radioactivity could be seen moving progressively down the axons. In the continuous infusion case, one would see a wave of radioactivity with a sharp but slowly spreading wavefront propagating at constant velocity down the axon. In the case of infusion for a few hours one would see at long times a slowly spreading pulse of radioactivity that looked normally distributed. It was to understand this behavior that Reed and Blum constructed PDE models for axonal transport [3], [33], [34]. These models did not have traveling wave solutions, but the data certainly looked like approximate traveling waves. In [35] it was shown by a perturbation theory argument that, in the asymptotic limit where the unbinding and binding rates  $k_2$  and  $k_1$  get large, the solution approaches a slowly spreading traveling wave or a normal pulse. Recently, in a series of papers, Friedman and coworkers have introduced new PDE models and proved these results rigorously [13], [14], [15], [16].

Probabilistic models for axonal transport were introduced and used for simulations already in the 1980s [38], [40]. However, rigorous work began with Lawler [27] in 1995 and was continued by Brooks [5], who used a continuous time stochastic model to show that the distribution of an individual particle is a spreading Gaussian at large times. Brooks also proved tail estimates for the central limit theorem and used them to estimate the error from normal. Independently, Bressloff [4] developed a discrete stepping model and performed an analysis under the assumption that the rate of unbinding and binding to transport is fast relative to lateral velocity over the length scale of interest. The author derived a characterization of the spreading wavefront of a particle entering at the nucleus and traveling to the distal end. This model served as the basis for later investigations by Newby and Bressloff [30], [31] wherein the authors characterize the axonal transport system as an intermittent search for hidden targets.

In this paper we revise the existing probabilistic models in order to study randomness in the system as a whole rather than exclusively from the point of view of an individual particle. Our goal is not only to recover and generalize previous results but also to investigate three specific, biologically important, properties of the whole stochastic system.

**1.1. Summary of results.** In section 2, we create a continuous-time Markov-chain queueing model for the axonal transport system. We show how to use experimental data to determine (or estimate) all the parameters of the model.

In section 3, we take the individual vesicle point of view. We prove the asymptotic forms in [35] with rigorous error estimates. We show that in the limit as the compartment size becomes small our model becomes the probabilistic model of [5]. We also show that in the limit as the lengths of the axon and time become large (with the scale of axon length on the order of the squared scale of time) our model becomes the PDE model of [34]. Since we assume that particles are independent, the

time evolution of the law of an individual will reflect the behavior of an ensemble of particles released at the same time.

In section 4 we adopt the full spatial system perspective to quantify stochasticity along the length of the axon. We begin by calculating in Proposition 4.1 the stationary distribution of a flow-through system that has sustained input from the nucleus, while particles are removed upon reaching the distal end. The stationary distribution has a product Poisson structure which allows for seamless transition between spatial scales.

With this mathematical model, we are able to make precise statements about three biologically important system properties. In the stationary distribution, the number of vesicles in each slice of the axon is independent and identically distributed (i.i.d.) (Proposition 4.1); however, this in and of itself is not sufficient to account for the sense that samples taken for different parts of the cell “look the same.” We compute in section 4.2 the coefficient of variation for the number of vesicles in sections of different length and show that the coefficient of variation is low for all but the smallest length scales. In section 4.3, we study the intermittent search problem posed by Newby and Bressloff [30] from the system point of view. Efficient transport to locations that need material must balance the speed of transport of material from the nucleus to the distal end of the cell with the rates of dissociation from the transport apparatus. We calculate the expected hitting time for a hidden target by all vesicles in the system. In so doing we encounter the counterintuitive result that while increasing the velocity of the motors while on transport increases the chance of any particular vesicle missing the target, the expected hitting time by the system actually decreases.

This hitting time approach is natural for needed material that is sparsely distributed throughout the axon, but when the needed cargo in question is more common, the time to replenishment is better addressed through the ODE approach that we develop in section 4.4. Due to the product structure of the law of the transient dynamics, this nonequilibrium behavior is determined by the  $2N$ -dimensional ODE governing the means. From this we estimate the timescale of return to equilibrium as a function of the length scale of interest.

**2. The model and its parameters.** Let  $L$  be the length of the axon, divided evenly into  $N = L/\delta$  lateral sections each of length  $\delta$ , equal to the step size of the motor protein. Within each section, we disregard any further spatial geometry and take the particles to be in one of two states:

- an on-transport state that steps laterally at a rate  $r = v/\delta$  per section, or
- an off-transport state that does not step laterally.

We use a  $2N$ -dimensional continuous-time Markov chain  $\{(Q_i(t), P_i(t)), i = 1, \dots, N\}$  to model the particle dynamics, where

- $Q_i(t)$  is the number of particles at time  $t$  in the on-transport state in section  $i$ ;
- $P_i(t)$  is the number of particles at time  $t$  in the off-transport state in section  $i$ .

DEFINITION 2.1 (stochastic compartmental model). *Let  $\{(Q_i(t), P_i(t)), i = 1, \dots, N\}_{t \geq 0}$  be a continuous-time Markov chain on the state space  $\mathbb{N}^N \times \mathbb{N}^N$  with the following transitions and time dependent rates:*

- (lateral transport)  $(Q_i, Q_{i+1}) \rightarrow (Q_i - 1, Q_{i+1} + 1)$  at rate  $rQ_i(t)$ ;
- (switch from on-transport to off-transport)  $(Q_i, P_i) \rightarrow (Q_i - 1, P_i + 1)$  at rate  $k_2Q_i(t)$ ;
- (switch from off-transport to on-transport)  $(P_i, Q_i) \rightarrow (P_i - 1, Q_i + 1)$  at rate  $k_1P_i(t)$ ;
- (production of new particles)  $(Q_1) \rightarrow (Q_1 + 1)$  at rate  $rq_0$ ;
- (removal of particles at distal end)  $(Q_N) \rightarrow (Q_N - 1)$  at rate  $rQ_N(t)$ .

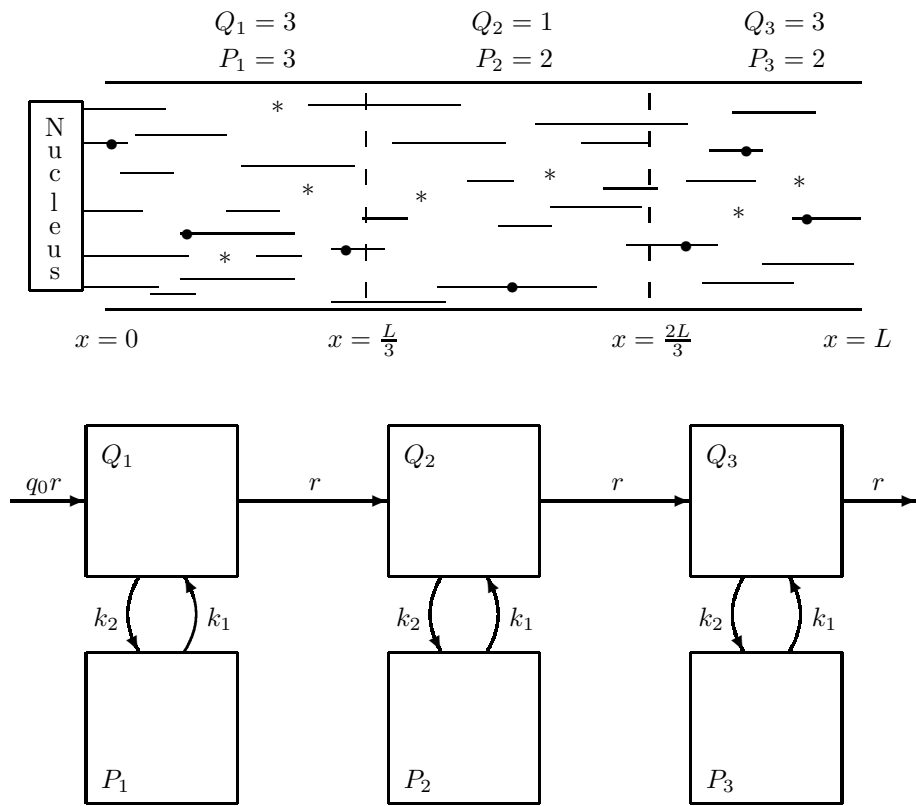


FIG. 1. A spatially discretized compartmental model for an on-and-off transport system.

The lateral transport rate,  $r = v/\delta$ , is inversely proportional to the length scale so that the mean number of particles per unit length is invariant with respect to rescaling  $\delta$ . We will assume that the rate of production  $q_0 = \delta\rho_0$  for some constant  $\rho_0 > 0$  in order for the mean number of particles in each compartment to scale with the size  $\delta$  of each compartment. This will imply that the mean number of particles per unit length scales as  $\rho_0$ . A graph of the model is depicted in Figure 1.

In order to ensure the Markov property, we use exponential random variables for the waiting times between transition events. Specifically, we mean that after a given event we assign a new independent random variable to each of the  $3N + 2$  possible next events, exponentially distributed with the appropriate rate parameter. The system of values updates according to the transition associated with the minimum of these waiting times. Then we create a new set of exponential random variables, and the process proceeds as before.

The advantage of computing explicit formulas for quantities that can be observed in experiments is that the experimental data can then be used to determine the parameter values in the model. For the characterization of the approximate wavefront speed and spreading in section 3 and the homogeneity calculations in section 4 we need order-of-magnitude estimates for the parameters. Actual parameter values will certainly differ depending on the particular neural tissue and the particular particles being transported. However, we can get order-of-magnitude estimates from existing data.

First we recall that fast transport has been observed to travel at speeds of 0.2 to 0.5  $m$  per day. We can assume that the average velocity of particles while physically bound to microtubules is roughly 1  $m$  per day, or  $v = 10^{-6}$   $m/s$ . We assume that the compartment size scales as the length scale of the individual steps of the motor protein, so  $\delta \sim 10^{-8}m$ . This implies that the rate parameter should be  $r = v/\delta = 100s^{-1}$ .

We now turn our attention to the on-off rates  $k_2$  and  $k_1$ . These can be determined from experimentally observed run lengths on the transport system. Indeed, Dixit et al. [9] show that a typical run along microtubules for dinein and kinesin is on the order of  $10^{-6}m$ . We can compare this with the theoretical run length of the model to determine off-rate  $k_2$ . Within the model, at each step on the transport mechanism the particle has a binary decision to jump laterally along the transport with probability  $r/(k_2 + r)$  or to jump off with probability  $k_2/(r + k_2)$ . The number of jumps along the transport system before jumping off is therefore geometrically distributed on the set  $\{0, 1, \dots\}$  with success probability  $\frac{r}{r+k_2}$ . It follows that the average number of steps in the run is  $\frac{r}{k_2}$  and therefore the average run length is  $\frac{r}{k_2} \times 10^{-8}m$ . Setting this equal to the average experimental run length of  $10^{-6}$  from [9], we see that  $\frac{r}{k_2} \sim 100$ , implying that  $k_2 \sim 1s^{-1}$ . As we will see in the computation of the stationary distribution in section 4.1, the ratio of the expected number of particles on the track to those off the track is  $\frac{k_2}{k_1}$ . Dixit et al. [9] found that approximately 75% of the particles were motile, so this ratio is approximately equal to 3. Since  $k_2 \sim 1s^{-1}$ , we see that  $k_1 \sim \frac{1}{3}$ .

It remains to estimate  $q_0$ . We will see in Proposition 4.1 that the mean number of particles per compartment is  $(1 + \frac{k_2}{k_1})q_0 = 4q_0$ . Of course, axons have a large variety of diameters, and larger axons will have more vesicles per unit length, so one expects a range of values for  $q_0$ . However, examination of a large number of electron micrographs of axonal cross-sections (see, for example, [19, Fig. 3], [20], [29]), which are typically 100  $nm$  thick, enables one to estimate the number of vesicles per 100  $nm$  segment. This number is typically in the range of 10 to 100, which implies that there are 1 to 10 vesicles per compartment in our model. Therefore,  $q_0$  is in the range 0.25 to 2.5 for various axons.

We remark that we are ignoring some aspects of the physics and the biology of axonal transport. We are not including diffusion of the vesicles off the track. We are treating the microtubule track as though it were a single continuous entity from one end of the axon to the other, when in fact it consists of numerous, separated, microtubule fragments. And, we are ignoring retrograde transport and the details of the motor proteins. Nevertheless, this simple model will enable us to investigate the homogeneity questions that are the main goal of this paper.

**3. Dynamics from the particle perspective.** In this section we calculate properties of the stochastic dynamics by using stochastic convergence theorems and stochastic averaging theorems from probability theory. We first see that, in the  $\delta \rightarrow 0$  limit, the law of the location of a single particle corresponds to that of a particle with a piecewise linear Markov motion. We then show this law can be approximated by the Green's function of a linear PDE. This enables us to obtain, as a special case, the asymptotic behavior of the PDE models for axonal transport in either a rate limiting or a perturbed setting.

**3.1. The active transport mode.** We first consider the simple case where the particle starts at  $X_0^\delta = 0$  and stays exclusively in active transport mode. Let  $X_t^\delta \in \{0, \delta, 2\delta, \dots, L\}$  be the lateral position of a particle at time  $t$ , and let  $n_t^\delta$  be the

number of jumps made by the particle as of time  $t$ . Observe that  $X^\delta = \delta n^\delta$ .

PROPOSITION 3.1. *Let  $k_2 = k_1 = 0$  and  $r = v/\delta > 0$ ; then the position of the particle satisfies, for any time  $t < \infty$ ,*

$$\sup_{s \leq t} |X_s^\delta - vs| \xrightarrow{\delta \rightarrow 0} 0 \quad a.s.$$

and, for  $B$  a standard Brownian motion,

$$\frac{1}{\sqrt{\delta}} (X_t^\delta - vt)_{t \geq 0} \xrightarrow{\delta \rightarrow 0} \sqrt{v}(B_t)_{t \geq 0}$$

in distribution on the Skorokhod space of cadlag (right continuous left limited) functions.

*Proof.* Since  $n^\delta$  is a Poisson process with rate  $r = v/\delta$ , defining  $N_t := n_{\delta t}^\delta$  we get a Poisson process  $N$  with rate  $v$ , and we have  $X_t^\delta = \delta N_{t/\delta}$ . Our results then follow directly from the functional law of large numbers (FLLN) and the functional central limit theorem (FCLT) for the Poisson process  $N$ .  $\square$

**3.2. The on/off dynamics.** We now consider a particle which undergoes transitions from the on-transport state to the off-transport state and back. Denote again by  $X_t^\delta \in \{0, \delta, 2\delta, \dots, L\}$  the lateral position of a particle at time  $t$ , and let  $n_t^\delta$  be the number of lateral transition jumps made by the particle as of time  $t$ . Observe that the particle will spend only a fraction of its time in active transport, and hence the lateral speed of the particle should be slower than before.

A noncompartmental stochastic model for axonal transport, introduced by Brooks in [5], is as follows. A particle can be in one of two states:

- an on-transport state with deterministic lateral velocity  $v$ , or
- an off-transport state with lateral velocity 0.

We use a two-dimensional Markov process to model the particle dynamics, where

- $X_t$  is the lateral position of this particle at time  $t$ , and
- $\xi_t$  is the indicator for whether it is on (1) or off (0) transport at time  $t$ .

DEFINITION 3.2 (stochastic noncompartmental model). *Let  $(X_t, \xi_t)_{t \geq 0}$  be a piecewise-linear Markov process with values in  $(\mathbf{R}_+, \{0, 1\})$  started at  $(X_0, \xi_0) = (0, 1)$  with the following dynamics:*

- (switch from on-transport to off-transport)  $(X_t, 1) \rightarrow (X_t, 0)$  at rate  $k_2$ ;
- (switch from off-transport to on-transport)  $(X_t, 0) \rightarrow (X_t, 1)$  at rate  $k_1$ ;
- (lateral travel)  $X_t = \int_0^t v \xi_s ds$ .

The path of  $(X_t)_{t \geq 0}$  consists of alternating sequences of Exponential( $k_2$ ) stretches of time where the lateral position increases linearly with speed  $v$ , and Exponential( $k_1$ ) stretches of time where it remains constant.

PROPOSITION 3.3. *Let  $k_2, k_1 > 0$  and  $r = v/\delta > 0$ ; then the position of the particle converges,*

$$(X_t^\delta)_{t \geq 0} \xrightarrow{\delta \rightarrow 0} (X_t)_{t \geq 0},$$

in distribution on the Skorokhod space of cadlag functions.

*Proof.* If we let  $\xi^\delta$  be the indicator of whether the particle in the compartmental model is on ( $\xi^\delta = 1$ ) or off ( $\xi^\delta = 0$ ) transport, then  $(X_t^\delta, \xi_t^\delta)_{t \geq 0}$  is a strong Markov process. We will see that  $\xi^\delta$  is a continuous-time Markov chain on  $\{0, 1\}$  (independent of  $\delta$ ), and conditionally on  $\xi^\delta$  the transition law of  $X$  is easily expressed. Likewise, for the noncompartmental model above,  $(X_t, \xi_t)_{t \geq 0}$  is a strong Markov process, with

$\xi$  the same continuous-time Markov chain on  $\{0, 1\}$  as  $\xi^\delta$ , and conditionally on  $\xi$  the change in  $X$  is easily given in terms of its linear speed and  $\xi$ .

We will start by showing that for any  $t > 0$ ,  $(X_t^\delta, \xi_t^\delta)$  converges to  $(X_t, \xi_t)$  in distribution as  $\delta \rightarrow 0$ . We then show that the finite dimensional distributions of  $(X^\delta, \xi^\delta)$  converge to those of  $(X, \xi)$ . A tightness argument finally implies (see Lemma 16.2 and Theorem 16.3 in [21]) that  $(X^\delta, \xi^\delta)$  converges to  $(X, \xi)$  in distribution on the Skorokhod space of cadlag processes.

Suppose that initially the particle is on transport at  $x_0$ , so  $(X_0^\delta, \xi_0^\delta) = (x_0, 1)$ . The first time  $\tau_1 = \inf\{t > 0 : \xi_t^\delta = 0\}$  at which the particle steps off transport has Exponential( $k_2$ ) distribution, irrespective of  $\delta$ . The first subsequent increment in time  $\sigma_1 = \inf\{t > 0 : \xi_{\tau_1+t}^\delta = 0\}$  after which the particle steps back on transport has Exponential( $k_1$ ) distribution, irrespective of  $\delta$  as well. This is repeated, and  $\xi^\delta$  is a simple continuous-time Markov chain on  $\{0, 1\}$  with transition rates  $k_2$  and  $k_1$ , from  $1 \rightarrow 0$  and  $0 \rightarrow 1$ , respectively.

Until time  $\tau_1$  the particle behaves as if it were in active transport ( $k_2 = k_1 = 0$ ) and conditionally on the value of  $\tau_1$ , for any  $0 \leq s \leq \tau_1$  the lateral change in position over time  $s$ ,  $X_s^\delta - X_0^\delta$ , is  $\delta n_s^\delta$ , where  $n_s^\delta$  is a Poisson( $rs$ ) random variable. Moreover, by results of Proposition 3.1, conditionally on the value of  $\tau_1$ , we have

$$\sup_{0 \leq s \leq \tau_1} |(X_s^\delta - X_0^\delta) - vs| \xrightarrow{\delta \rightarrow 0} 0 \quad \text{a.s.}$$

To get the unconditioned law of  $X_{\tau_1}^\delta - X_0^\delta$ , we observe that the number of boxes the particle traverses before it steps off  $n_{\tau_1}^\delta$  has a Geometric( $k_2/(k_2+r)$ ) distribution (note that a Poisson rate  $r$  process sampled at an Exponential( $k_2$ ) time independent of the process has this distribution). Since  $k_2/\delta(k_2+r) \rightarrow k_2/v$  as  $\delta \rightarrow 0$ ,  $X_{\tau_1}^\delta - X_0^\delta = \delta n_{\tau_1}^\delta$  converges in distribution to Exponential( $k_2/v$ ) variable.

Consider now the particle in the noncompartmental model. It is immediate from the definition of the model that  $\xi$  has the same law of a continuous-time Markov chain on  $\{0, 1\}$  with transition rates  $k_2$  and  $k_1$ , from  $1 \rightarrow 0$  and  $0 \rightarrow 1$ , as  $\xi^\delta$ . Since we prove our convergence in law results by first conditioning on the values of  $\xi^\delta$  and  $\xi$ , we can without loss of generality henceforth assume  $\xi^\delta = \xi$  a.s. and drop its superscript.

Suppose that initially the particle in the noncompartmental model is on transport at  $x_0$ , so  $(X_0, \xi_0) = (x_0, 1)$ . Conditionally on  $\tau_1$ , for all  $0 \leq s \leq \tau_1$ ,  $X_s - X_0 = vs$  and  $X_{\tau_1} - X_0 = v\tau_1$ ; hence, unconditionally,  $X_{\tau_1} - X_0$  is an Exponential( $k_2/v$ ) random variable. We now have both  $\sup_{0 \leq s \leq \tau_1} |(X_s^\delta - X_0^\delta) - (X_s - X_0)| \rightarrow 0$  a.s. and  $X_{\tau_1}^\delta - X_0^\delta \Rightarrow X_{\tau_1} - X_0$ .

Between times  $\tau_1$  and  $\sigma_1$  the particle in both models stays in place, so conditionally on values of  $\tau_1, \sigma_1$ , and of  $X_{\tau_1}^\delta, X_{\tau_1}$ ,  $\sup_{\tau_1 \leq s \leq \tau_1 + \sigma_1} |(X_s^\delta - X_{\tau_1}^\delta) - (X_s - X_{\tau_1})| \equiv 0$ , and  $X_{\tau_1 + \sigma_1}^\delta - X_0^\delta \Rightarrow X_{\tau_1 + \sigma_1} - X_0$ . At time  $\tau_1 + \sigma_1$ , the same process starts over from initial values  $(X_{\tau_1}^\delta, 1)$  and  $(X_{\tau_1}, 1)$  in the compartmental and noncompartmental models, respectively.

Let  $\tau_1, \sigma_1, \tau_2, \dots$  be the sequence of time increments between the consecutive times when the particle in both models gets off and gets back on transport, let  $\sigma_0 = 0$ , and for  $i \geq 1$

$$\tau_i = \inf\{t > 0 : \xi_{\sigma_{i-1}+t} = 0\}, \quad \sigma_i = \inf\{t > 0 : \xi_{\tau_i+t} = 1\}.$$

Then  $(\tau_i)_{i \geq 1}$  and  $(\sigma_i)_{i \geq 1}$  are independent sequences of i.i.d. Exponential( $k_2$ ) and Exponential( $k_1$ ) variables, respectively. For any  $t > 0$ , let  $\eta_t$  be the number of times

the particle in either model gets back on transport until time  $t$  and  $\eta'_t$  the number of times it gets off:

$$\eta_t = \inf \left\{ k \geq 0 : \sum_{i=1}^k (\tau_i + \sigma_i) \leq t \right\}, \quad \eta'_t = \inf \left\{ k' \geq 0 : \sum_{i=1}^{k'} \tau_i + \sum_{i=1}^{k'-1} \sigma_i \leq t \right\}.$$

Note that  $\eta'_t = \eta_t$  iff  $\xi_t = 1$ , and  $\eta'_t = \eta_t + 1$  iff  $\xi_t = 0$ . Let  $\tilde{\tau}_t$  be the last time before time  $t$  that the particle changed whether it was on or off transport; that is,  $\tau_t = \sup\{0 \leq s \leq t : \xi_{s-} \neq \xi_s\}$ . Then, we have

$$\tau_t = \begin{cases} \sum_{i=1}^{\eta_t} (\tau_i + \sigma_i) & \text{if } \eta'_t = \eta_t, \\ \sum_{i=1}^{\eta'_t} \tau_i + \sum_{i=1}^{\eta_t} \sigma_i & \text{if } \eta'_t = \eta_t + 1. \end{cases}$$

If  $\eta'_t = \eta_t$ , then from time  $\tau_t$  to  $t$  the particle is in active transport, and the same convergence argument as before implies that conditionally on the values of  $\tau_t$  and  $X_{\tau_t}^\delta$ ,

$$\sup_{\tau_t \leq s \leq t} |(X_s^\delta - X_{\tau_t}^\delta) - v(s - \tau_t)| \rightarrow 0 \text{ a.s.}$$

Also  $X_{\tau_t}^\delta - X_0^\delta = \delta n_{\tau_t}^\delta$  where  $n_{\tau_t}^\delta$  is the number of boxes the particle traverses by time  $t$ . Conditionally on the value of  $\eta_t$ ,  $n_{\tau_t}^\delta$  is a sum of  $\eta_t$  i.i.d. Geometric( $k_2/(k_2 + r)$ ) random variables; hence  $X_{\tau_t}^\delta - X_0^\delta$  converges in distribution to a sum of  $\eta_t$  i.i.d. Exponential( $k_2/v$ ) random variables. In the noncompartmental model, if  $\eta'_t = \eta_t$ , then conditionally on the values of  $\eta_t$  and  $\tau_t$ , for  $\tau_t \leq s \leq t$ ,  $X_s - X_{\tau_t} = v(s - \tau_t)$  and conditionally only on the value of  $\eta_t$ ,  $X_{\tau_t}^\delta - X_0$  is a sum of  $\eta_t$  i.i.d. Exponential( $k_2/v$ ) variables. Hence, conditionally on  $\eta_t$  and  $\tau_t$ ,  $\sup_{\tau_t \leq s \leq t} |(X_s^\delta - X_{\tau_t}^\delta) - (X_s - X_{\tau_t})| \rightarrow 0$  a.s., and conditionally only on  $\eta_t$ ,  $X_{\tau_t}^\delta - X_0^\delta \Rightarrow X_{\tau_t} - X_0$ .

If  $\eta'_t = \eta_t + 1$ , then from time  $\tau_t$  to  $t$  the particle in both models stays in place, so conditionally on  $\eta_t$  and  $\tau_t$ ,  $\sup_{\tau_t \leq s \leq t} |(X_s^\delta - X_{\tau_t}^\delta) - (X_s - X_{\tau_t})| = 0$  a.s. Also, conditionally only on the value of  $\eta'_t$ ,  $n_{\tau_t}^\delta$  is a sum of  $\eta'_t$  i.i.d. Geometric( $k_2/(k_2 + r)$ ) variables; hence  $X_{\tau_t}^\delta - X_0^\delta$  converges in distribution to a sum of  $\eta'_t$  i.i.d. Exponential( $k_2/v$ ) variables. In the noncompartmental model, if  $\eta'_t = \eta_t + 1$ , conditionally only on the value of  $\eta'_t$ ,  $X_{\tau_t}^\delta$  is a sum of  $\eta'_t$  i.i.d. Exponential( $k_2/v$ ) variables. Hence, conditionally only on  $\eta'_t$ ,  $X_{\tau_t}^\delta - X_0^\delta \Rightarrow X_{\tau_t} - X_0$ .

Now, integrating over the possible values of  $\eta_t$ ,  $\eta'_t$ , and  $\tau_t$ , we get that for any  $t \geq 0$ ,  $(X_t^\delta, \xi_t^\delta) \Rightarrow (X_t, \xi_t)$ . Convergence of finite dimensional distributions follows from an iterative use of the Markov property of  $(X^\delta, \xi^\delta)$  and  $(X, \xi)$ , and the fact that the increments of both  $(X^\delta, \xi^\delta)$  and  $(X, \xi)$  are stationary.

In order to verify tightness (see Theorem 16.11 in [21]) of the sequence of Markov processes  $\{(X^\delta, \xi^\delta)\}_{\delta > 0}$ , because  $(X^\delta, \xi^\delta)$  has stationary increments and is strong Markov, it will suffice to check that for any  $\epsilon > 0$

$$\lim_{h \rightarrow 0} \limsup_{\delta \rightarrow 0} \mathbb{P}\{ \|(X_h^\delta, \xi_h^\delta) - (X_0^\delta, \xi_0^\delta)\| > \epsilon \} = 0,$$

where  $\|(x_1, \xi_1) - (x_2, \xi_2)\| = |x_1 - x_2| + |\xi_1 - \xi_2|$  is a distance metric on  $\mathbf{R}_+ \times \{0, 1\}$ . For any  $\delta > 0$ , the first change in the continuous-time Markov chain  $\xi^\delta$  happens after an Exponential( $k$ ) time (where  $k = k_2$  or  $k = k_1$  depending on whether  $\xi_0^\delta = 1$  or  $\xi_0^\delta = 0$ ) and is independent of  $\delta$ . Hence, at time  $h$  later,  $\mathbb{P}\{\xi_h^\delta \neq \xi_0^\delta\} \leq 1 - e^{-kh}$ . Moreover,



irrespective of the value of  $\xi^\delta$ , at time  $h$  later the value of  $|X_h^\delta - X_0^\delta| \leq \delta n_h^\delta$ , where  $n^\delta$  is a Poisson process with rate  $r = v/\delta$ . Hence,  $\mathbb{P}\{|X_h^\delta - X_0^\delta| > \epsilon\} \leq \delta \mathbb{E}[n_h^\delta]/\epsilon = vh/\epsilon$ . Combining the two gives

$$\mathbb{P}\{||(X_h^\delta, \xi_h^\delta) - (X_0^\delta, \xi_0^\delta)|| > \epsilon\} \leq 1 - e^{-kh} + vh/\epsilon \text{ for any } \delta > 0,$$

and the desired limit follows.  $\square$

The process  $(X^\delta, \xi^\delta) : t \in [0, \infty) \mapsto (X_t^\delta, \xi_t^\delta) \in \delta\mathbb{Z}_+ \times \{0, 1\}$  is a Markov process with cadlag paths whose generator is given by

$$\begin{aligned} A^\delta f(x, \xi) &= r\xi[f(x + \delta, \xi) - f(x, \xi)] \\ &\quad + k_2\xi[f(x, \xi - 1) - f(x, \xi)] + k_1(1 - \xi)[f(x, \xi + 1) - f(x, \xi)] \end{aligned}$$

for all  $f \in \mathcal{D}(A^\delta) = \mathcal{C}^0(\delta\mathbb{Z}_+ \times \{0, 1\})$ .

The piecewise-linear process  $(X, \xi) : t \in [0, \infty) \mapsto (X_t, \xi_t) \in \mathbb{R}_+ \times \{0, 1\}$  is a Markov process with continuous paths whose generator is the closure of the operator

$$\begin{aligned} Af(x, \xi) &= v\xi\partial_x f(x, \xi) \\ &\quad + k_2\xi[f(x, \xi - 1) - f(x, \xi)] + k_1(1 - \xi)[f(x, \xi + 1) - f(x, \xi)] \end{aligned}$$

for all  $f \in \mathcal{D}(A^\delta) = \mathcal{C}^{1,0}(\mathbb{R}_+ \times \{0, 1\})$ .

Letting  $\iota^\delta : \delta\mathbb{Z}_+ \times \{0, 1\} \mapsto \mathbb{R}_+ \times \{0, 1\}$  be an embedding and  $f^\delta = f \circ \iota^\delta$ , then  $A^\delta f^\delta \rightarrow Af$  as  $\delta \rightarrow 0$  for all  $f \in \mathcal{C}^{1,0}(\mathbb{R}_+ \times \{0, 1\})$  implies that the finite dimensional distributions of  $(X^\delta, \xi^\delta)$  converge to those of  $(X, \xi)$ . Verification of additional conditions (see Theorem 19.25 of [21]) would also imply convergence of processes with generators  $\{(A^\delta, \mathcal{D}(A^\delta))\}_{\delta>0}$  to the process with generator  $(A, \mathcal{D}(A^\delta))$ ; however, we thought this way of showing convergence in law was not as instructive.

The fact that an individual particle will have the distribution given by Proposition 3.3 as the size of the boxes decreases means that our model is a microscopic version of the stochastic model used by Brooks [5] and that the hydrodynamic limit of our model as  $\delta \rightarrow 0$  is equal to the macroscopic stochastic model from [5].

An approximation of the particle's position  $X_t$  is obtained in [5] to be  $X_t \approx \mu t + \sqrt{t}\sigma Z$  as  $t \rightarrow \infty$ , where  $\mu = k_1v/(k_2 + k_1)$ ,  $\sigma = 2k_2k_1v^2/(k_2 + k_1)^3$ , and  $Z$  is a standard normal variable. That approximation is valid only for large fixed values of  $t$ , while we next extend this result to give an approximation for the whole time trajectory of the particle's path. This is accomplished by the following FCLT for the position of the particle undergoing stochastic transport.

PROPOSITION 3.4. *Let  $X$  be the position of a particle following the piecewise-linear Markov process from Proposition 3.3 started at  $X_0 = 0$  on transport; then*

$$\sup_{s \leq t} \left| \frac{X_{ns}}{n} - \frac{k_1}{k_2 + k_1}vs \right| \xrightarrow[n \rightarrow \infty]{} 0 \quad a.s. \quad \forall t > 0,$$

and, if  $B$  denotes a standard Brownian motion,

$$\sqrt{n} \left( \frac{X_{nt}}{n} - \frac{k_1v}{k_2 + k_1}t \right)_{t \geq 0} \xrightarrow[n \rightarrow \infty]{\Rightarrow} \sqrt{\frac{2k_2k_1v^2}{(k_2 + k_1)^3}} (B_t)_{t \geq 0}$$

in distribution on the space of continuous functions.

*Proof.* For these results we use the notion of stochastic averaging [22], [23], [25]. Note that the indicator process  $\xi$  for being on- or off-transport is independent of the

position  $X$  of the particle. Hence, the position of the particle  $X$  is a linear random evolution process [10, Chap. 12], [18] driven by the independent indicator process  $\xi$ . The generator of  $(X, \xi)$  is the closure of the operator

$$Af(x, \xi) = \sigma(\xi)\partial_x f(x, \xi) + \lambda(\xi)[f(x, s(\xi)) - f(x, \xi)]$$

for all  $f \in \mathcal{D}(A) = \mathcal{C}_0^{1,0}(\mathbb{R}_+ \times \{0, 1\})$  (the space of all continuously differentiable functions in  $x$  continuous in  $\xi$  and vanishing at infinity), where

$$\sigma(\xi) = v\xi, \quad \lambda(\xi) = k_2\xi + k_1(1 - \xi), \quad \text{and} \quad s(\xi) = 1 - \xi$$

(when  $\xi = 1$ ,  $\sigma = v$ ,  $\lambda = k_2$ , and  $s = 0$ , and when  $\xi = 0$ ,  $\sigma = 0$ ,  $\lambda = k_1$ , and  $s = 1$ ). In other words, if  $(X_0, \xi_0) = (0, 1)$ , we have

$$X_t = \int_0^t v\xi_s ds, \quad \xi_t = \frac{1}{2}\left(1 + (-1)^{Y(\int_0^t \lambda(\xi_s) ds)}\right),$$

where  $Y$  is a rate 1 Poisson process, and  $Y(\int_0^t \lambda(\xi_s) ds)$  is a counting process of the number of switches of  $\xi$  until time  $t$ .

Rescaling time and position of the process by  $1/n$ , we get that  $(\frac{X_n}{n}, \xi_n)$  satisfies

$$\frac{X_{nt}}{n} = \int_0^t v\xi_{ns} ds, \quad \xi_{nt} = \frac{1}{2}\left(1 + (-1)^{Y(n \int_0^t \lambda(\xi_{ns}) ds)}\right),$$

and its generator is

$$A^n f(x, \xi) = v\xi\partial_x f(x, \xi) + n(k_2\xi + (k_1(1 - \xi)))[f(x, s(\xi)) - f(x, \xi)]$$

for all  $f \in \mathcal{D}(A^n) = \mathcal{C}_0^{1,0}(\mathbb{R}_+ \times \{0, 1\})$ . Note that  $\xi_n$  switches at rate proportional to  $n$ , forming an ergodic Markov chain with stationary distribution  $\pi(1) = k_1/(k_2 + k_1)$ ,  $\pi(0) = k_2/(k_2 + k_1)$ , and  $\int \sigma(\xi)\pi(\xi) = v \int \xi\pi(\xi) = vk_1/(k_2 + k_1)$ . Hence, the strong ergodic theorem implies that

$$\frac{X_n}{n} = \frac{1}{n} \int_0^n v\xi_s ds \xrightarrow[n \rightarrow \infty]{} \frac{vk_1}{k_2 + k_1} \quad \text{a.s.}$$

We can extend this to a functional statement on any finite time interval  $[0, t]$ . Fix  $t > 0$  and take any  $\Delta > 0$ ; then there exists  $n_\Delta < \infty$  a.s. such that

$$\left| \frac{X_n}{n} - \frac{vk_1}{k_2 + k_1} \right| < \frac{\Delta}{t} \quad \forall n > n_\Delta.$$

Now, let  $M = \sup_{n \leq n_\Delta} |X_n - k_1/(k_2 + k_1)vn|$ , which is finite a.s. since  $n_\Delta < \infty$  a.s. Let  $n > M/\Delta$ . Then, for any  $0 \leq s \leq t$  we have that either  $ns > n_\Delta$ , in which case

$$\left| \frac{X_{ns}}{n} - \frac{vk_1}{k_2 + k_1}s \right| = \left| \left( \frac{X_{ns}}{ns} - \frac{vk_1}{k_2 + k_1} \right) s \right| < \frac{\Delta}{t}s \leq \Delta,$$

or  $ns \leq n_\Delta$ , in which case

$$\left| \frac{X_{ns}}{n} - \frac{vk_1}{k_2 + k_1}s \right| = \frac{1}{n} \left| X_{ns} - \frac{vk_1}{k_2 + k_1}ns \right| < \frac{M}{n} < \Delta,$$

implying that we have  $\sup_{0 \leq s \leq t} \left| \frac{1}{n}X_{ns} - vk_1/(k_2 + k_1)s \right| < \Delta$  whenever  $n > M/\Delta$ .

Once we rescale the position for the particle by  $1/\sqrt{n}$  and time by  $1/n$ ,  $\xi$  still changes at a much faster rate than the position of the particle  $X$ . The generator of the rescaled centered process  $(X^n, \xi^n)$  defined as

$$X_t^n := \sqrt{n} \left( \frac{X_{nt}}{n} - \frac{k_1 v}{k_2 + k_1} t \right), \quad \xi_t^n := \xi_{nt}$$

is the closure of the operator

$$\bar{A}^n f(x, \xi) = \left( \sqrt{n} \sigma(\xi) - \frac{k_1 v}{k_2 + k_1} \right) \partial_x f(x, \xi) + n \lambda(\xi) (f(x, s(\xi)) - f(x, \xi))$$

on  $f \in \mathcal{D}(A^n) = \mathcal{C}_0^{1,0}(\mathbb{R} \times \{0, 1\})$ . We will use the stochastic averaging theorem (Theorem 2.1 of [25]) to show that the paths of the centered rescaled process converge in distribution to paths of a Brownian motion with a diffusion coefficient equal to  $2k_2 k_1 / (k_2 + k_1)^3 v^2$ .

Let  $h(\xi)$  be the function

$$h(\xi) = v \frac{k_2 k_1}{(k_2 + k_1)^3} \frac{1}{\lambda(\xi)} = v \frac{k_2 k_1}{(k_2 + k_1)^3} \frac{1}{k_2 \xi + k_1 (1 - \xi)}.$$

Then  $h(1) = vk_1 / (k_2 + k_1)^2$ ,  $h(0) = -vk_2 / (k_2 + k_1)^2$  imply that  $h(s(1)) - h(1) = -v / (k_2 + k_1)$ ,  $h(s(0)) - h(0) = v / (k_2 + k_1)$ , which in turn imply that  $\lambda(1)(h(s(1)) - h(1)) = -vk_2 / (k_2 + k_1)$ ,  $\lambda(0)(h(s(0)) - h(0)) = vk_1 / (k_2 + k_1)$ , so that for  $\xi \in \{0, 1\}$

$$\lambda(\xi)(h(s(\xi)) - h(\xi)) = - \left( v\xi - v \frac{k_1}{k_2 + k_1} \right).$$

Now, for any  $f \in \mathcal{C}_0^2(\mathbb{R})$  define a sequence of functions  $f^n \in \mathcal{C}_0^{1,0}(\mathbb{R} \times \{0, 1\})$  by

$$f^n(x, \xi) = f(x) + \frac{1}{\sqrt{n}} h(\xi) \partial_x f(x).$$

Then  $f^n \rightarrow f$  as  $n \rightarrow \infty$  and

$$\begin{aligned} \bar{A}^n f^n(x, \xi) &= \sqrt{n} \left( v\xi - v \frac{k_1}{k_2 + k_1} \right) \partial_x f(x) + \left( v\xi - v \frac{k_1}{k_2 + k_1} \right) h(\xi) \partial_x^2 f(x) \\ &\quad + \sqrt{n} \lambda(\xi) (h(s(\xi)) - h(\xi)) \partial_x f(x) \\ &= \left( v\xi - v \frac{k_1}{k_2 + k_1} \right) h(\xi) \partial_x^2 f(x) = \bar{A} f(x), \end{aligned}$$

where  $\bar{A}$  is defined on  $\mathcal{D}(\bar{A}) = \mathcal{C}_0^2(\mathbb{R})$  by

$$\bar{A} f(x) = \frac{k_2 k_1 v^2}{(k_2 + k_1)^3} \partial_x^2 f(x).$$

Define a sequence of processes

$$\varepsilon_t^{f,n} = \frac{1}{\sqrt{n}} h(\xi_t^n) \partial_x f(X_t^n) = f^n(X_t^n, \xi_t^n) - f(X_t^n).$$

Then our earlier calculation implies that for any  $f \in \mathcal{D}(\bar{A})$

$$f(X_t^n) - \int_0^t \bar{A} f(X_s^n) ds + \varepsilon_t^{f,n} = f^n(X_t^n, \xi_t^n) - \int_0^t A^n f(X_s^n, \xi_s^n) ds$$

is a sequence of martingales. Since  $f \in C_0^2(\mathbb{R})$ ,  $\xi_t^n \in \{0, 1\}$ , it is clear that

$$\sup_n \mathbb{E} \left[ \int_0^t |\bar{A}f(X_s^n)|^2 ds \right] < \infty \quad \text{and} \quad \mathbb{E} \left[ \sup_{s \leq t} |\varepsilon_s^{f,n}| \right] \xrightarrow{n \rightarrow \infty} 0.$$

In order to apply Theorem 2.1 of [25] on stochastic averaging it is left only to show that the process  $X^n$  satisfies the compact containment condition; that is, for any  $t > 0$  and  $\Delta > 0$  there exists a compact set  $K \subset \mathbb{R}$  such that

$$\inf_n \mathbb{P}\{X_s^n \in K \forall s \leq t\} \geq 1 - \Delta.$$

This follows from the fact that  $X_t^n + h(\xi_t^n)/\sqrt{n}$  is a sequence of martingales (let  $f(x) = x$ ) with mean

$$\mathbb{E} \left[ X_t^n + \frac{h(\xi_t^n)}{\sqrt{n}} \right] = X_0^n + \frac{h(\xi_0^n)}{\sqrt{n}} = \frac{k_2 k_1}{(k_2 + k_1)^3} \frac{v^2}{\sqrt{n}}$$

and second moment (let  $f(x) = x^2$ )

$$\mathbb{E} \left[ \left( X_t^n + \frac{h(\xi_t^n)}{\sqrt{n}} \right)^2 \right] = \frac{k_2 k_1}{(k_2 + k_1)^3} v^2 2t + \frac{\mathbb{E}[h(\xi_t^n)]}{n}.$$

So, by Doob’s inequality,

$$\begin{aligned} \mathbb{P} \left\{ \sup_{s \leq t} \left| X_s^n + \frac{h(\xi_s^n)}{\sqrt{n}} \right| \geq C \right\} &\leq \frac{4}{C^2} \mathbb{E} \left[ \left( X_t^n + \frac{h(\xi_t^n)}{\sqrt{n}} \right)^2 \right] \\ &= \frac{4}{C^2} \left( \frac{k_2 k_1}{(k_2 + k_1)^3} v^2 2t + \frac{\mathbb{E}[h(\xi_t^n)]}{n} \right). \end{aligned}$$

Noting that  $h_{\min} := v \min(k_2, k_1)/(k_2 + k_1)^2 \leq h \leq h_{\max} := v \max(k_2, k_1)/(k_2 + k_1)^2$  and choosing  $C$  (given on  $t$  and  $\Delta$ ) so that the right-hand side of the inequality with  $n = 1$  is less than  $\Delta$  shows that with  $K = [-C - h_{\max}, C - h_{\min}]$  the compact containment condition holds for  $(X^n)_{n \geq 1}$ .

Now Theorem 2.1 of [25] implies that  $X^n \Rightarrow W$  in distribution on the Skorokhod space of continuous functions, where  $W$  is a process with generator  $\bar{A}$  and consequently has the same distribution as  $\sqrt{2k_2 k_1 v^2 / (k_2 + k_1)^3} B$ , where  $B$  is a standard Brownian motion.  $\square$

**3.3. Connection to PDE models.** In order to demonstrate the connection between our model and the PDEs seen in [35], [34], [15], consider the process  $(X, \xi)$  of the particle following the piecewise-linear Markov process, and for any  $x \geq 0$ ,  $t \geq 0$  let

$$q(x, t) = \mathbb{P}\{X_t \in dx, \xi_t = 1\}/dx, \quad p(x, t) = \mathbb{P}\{X_t \in dx, \xi_t = 0\}/dx$$

denote the probability densities of the particle’s location  $x$  on- and off-transport, respectively, over time. Kolmogorov forward equations for  $(X, \xi)$  imply that  $q$  and  $p$  satisfy the system of PDEs

$$(3.1) \quad \partial_t q(x, t) + v \partial_x q(x, t) = -k_2 q(x, t) + k_1 p(x, t),$$

$$(3.2) \quad \partial_t p(x, t) = k_2 q(x, t) - k_1 p(x, t).$$

When  $k_2 = 0 = k_1$ , the limiting PDE is simple linear transport:  $(\partial_t + v\partial_x)q(x, t) = 0$ . The initial condition  $q(x, 0) = \delta_0(x)$  corresponds to the density of a single particle at the origin at  $t = 0$ . The time evolution via simple linear transport is a translation of the delta function, while the time evolution via (3.1) and (3.2) will have a spreading profile. This is clear from the macroscopic limits of  $(X^\delta, \xi^\delta)$  as  $\delta \rightarrow 0$ . When  $k_2 = k_1 = 0$ , the particle never switches off from traveling on transport at speed  $v$  and is deterministic, as seen in Proposition 3.1. When  $k_2, k_1 > 0$  the particle follows a truly stochastic process  $(X, \xi)$  with a nonzero variance, as seen in Propositions 3.3 and 3.4.

In the experiments described in the introduction one sees “approximate” traveling waves of radioactivity in the axons in the sense that there is a slowly spreading wavefront moving at constant velocity away from the cell body. Equations (3.1) and (3.2) are linear and do not have solutions that are bounded traveling waves. It was shown by a perturbation theory argument in [35], [34] that as  $\varepsilon \rightarrow 0$  the solutions of

$$(3.3) \quad \varepsilon(\partial_t + v\partial_x)q^\varepsilon(x, t) = -k_2q^\varepsilon(x, t) + k_1p^\varepsilon(x, t),$$

$$(3.4) \quad \varepsilon\partial_t p^\varepsilon(x, t) = k_2q^\varepsilon(x, t) - k_1p^\varepsilon(x, t),$$

subject to  $q^\varepsilon(0, t) = q_0$ , are to leading order

$$q^\varepsilon(x, t) = c_1 H\left(\frac{x - \mu t}{\varepsilon^{1/2}}, t\right), \quad p^\varepsilon(x, t) = c_2 H\left(\frac{x - \mu t}{\varepsilon^{1/2}}, t\right),$$

where  $H$  satisfies the heat equation

$$(3.5) \quad \partial_s H(y, s) = \frac{\sigma^2}{2} \partial_{yy} H(y, s), \quad H(y, 0) = \chi_{(-\infty, 0)},$$

$$\mu = \frac{k_2 v}{k_2 + k_1}, \quad \sigma^2 = \frac{2k_2 k_1 v^2}{(k_2 + k_1)^3}, \quad c_1 = \frac{k_1}{k_2 + k_1}, \quad c_2 = \frac{k_2}{k_2 + k_1}.$$

This asymptotic form is valid for small  $\varepsilon$ , that is, for large  $k_2$  and  $k_1$ . However, if we set  $q(x, t) = q^\varepsilon(\frac{x}{\varepsilon}, \frac{t}{\varepsilon})$  and  $p(x, t) = p^\varepsilon(\frac{x}{\varepsilon}, \frac{t}{\varepsilon})$ , then  $q$  and  $p$  satisfy (3.3) and (3.4), so the solutions of (3.1) and (3.2) behave like approximate traveling waves for large  $t$  and large  $x$  whether or not  $k_2$  and  $k_1$  are large. These results have been proven rigorously by Friedman and coworkers [13], [14], [15], [16].

To see that our Proposition 3.4 provides another rigorous proof of these properties, albeit using stochastic methods, consider the process  $(\varepsilon X_{t/\varepsilon}, \xi_{t/\varepsilon})$  (that is,  $(\frac{1}{n} X_n, \xi_n)$  with  $n = 1/\varepsilon$ ), and let

$$q^\varepsilon(x, t) = \mathbb{P}\{\varepsilon X_{t/\varepsilon} \in dx, \xi_{t/\varepsilon} = 1\}/dx, \quad p^\varepsilon(x, t) = \mathbb{P}\{\varepsilon X_{t/\varepsilon} \in dx, \xi_{t/\varepsilon} = 0\}/dx$$

be the probability densities for this process. The generator of this process is  $A^n$  ( $n = 1/\varepsilon$ ), so the Kolmogorov forward equations imply that  $q^\varepsilon$  and  $p^\varepsilon$  satisfy the system of PDEs (3.3) and (3.4). Our result from Proposition 3.4 states that

$$\mathbb{P}\{\varepsilon X_{t/\varepsilon} \in dx\}/dx \approx H\left(\frac{x - \mu t}{\varepsilon^{1/2}}, t\right)$$

for small  $\varepsilon > 0$ , where  $H$  satisfies (3.5). Hence,  $q^\varepsilon + p^\varepsilon \approx H(\frac{x - \mu t}{\varepsilon^{1/2}}, t)$ , and  $\mathbb{P}\{\xi_{t/\varepsilon} = 1\} \approx k_1/(k_2 + k_1)$ ,  $\mathbb{P}\{\xi_{t/\varepsilon} = 0\} \approx k_2/(k_2 + k_1)$  gives the result that  $q^\varepsilon(x, t)$  and  $p^\varepsilon(x, t)$  are well approximated by  $\frac{k_1}{k_2 + k_1} H(\frac{x - \mu t}{\varepsilon^{1/2}}, t)$  and  $\frac{k_2}{k_2 + k_1} H(\frac{x - \mu t}{\varepsilon^{1/2}}, t)$ .

#### 4. Dynamics from the spatial system perspective.

**4.1. The spatial system in equilibrium.** We are now ready to characterize the steady state dynamics induced by continually adding particles from the nucleus and removing them when they reach the distal end of the cell.

**PROPOSITION 4.1.** *Let  $\{(Q_i(t), P_i(t)), i = 1, \dots, N\}_{t \geq 0}$  be the number of particles in the axonal transport system with compartments of size  $\delta$ , on- and off-transport, respectively. Suppose the rate of production of particles from the source is  $r q_0 = v \rho_0$ . Then this Markov chain has the product-form stationary distribution*

$$(Q_i, P_i) \sim \text{Pois}(q_0) \otimes \text{Pois}\left(\frac{k_2 q_0}{k_1}\right),$$

where all  $\{(Q_i, P_i), i = 1, \dots, N\}$  are mutually independent.

*Proof.* Since the production rate is  $r q_0$ , the generator of the process  $(Q_1, P_1)$  is

$$\begin{aligned} \mathcal{A}_{q_0} f(q, p) &= [f(q+1, p) - f(q, p)] r q_0 + [f(q-1, p) - f(q, p)] r q \\ &\quad + [f(q-1, p+1) - f(q, p)] k_2 q + [f(q+1, p-1) - f(q, p)] k_1 p. \end{aligned}$$

If we use  $f(q, p) = Q_1(t)$  and  $f(q, p) = P_1(t)$  and take expectations, we get a system of ODEs governing the change in  $\mathbb{E}[Q_1], \mathbb{E}[P_1]$  over time,

$$\begin{aligned} \frac{d\mathbb{E}[Q_1](t)}{dt} &= r q_0 - r \mathbb{E}[Q_1(t)] - k_2 \mathbb{E}[Q_1(t)] + k_1 \mathbb{E}[P_1(t)], \\ \frac{d\mathbb{E}[P_1](t)}{dt} &= k_2 \mathbb{E}[Q_1(t)] - k_1 \mathbb{E}[P_1(t)], \end{aligned}$$

indicating that in equilibrium in the first section the mean numbers of on-transport particles and off-transport particles are  $\mathbb{E}[Q_1] = q_0$  and  $\mathbb{E}[P_1] = \mathbb{E}[Q_1] \frac{k_2}{k_1} = q_0 k_2 / k_1$ , respectively. Let  $\pi_{q_0}(q, p) = \pi_{\lambda_Q}(q) \otimes \pi_{\lambda_P}(p)$  be a product of two independent Poisson distributions with rates  $\lambda_Q = q_0$  and  $\lambda_P = q_0 k_2 / k_1$ , respectively. To show that  $\pi_{q_0}(q, p)$  is a stationary distribution for the process  $(Q_1, P_1)$ , we need to check that

$$\sum_{q=0}^{\infty} \sum_{p=0}^{\infty} \mathcal{A}_{q_0} f(q, p) \pi_{q_0}(q, p) = 0$$

for any choice of function  $f \in \mathcal{D}(\mathcal{A}_{q_0})$ :

$$\begin{aligned} &\sum_{q=0}^{\infty} \sum_{p=0}^{\infty} \mathcal{A}_{q_0} f(q, p) e^{-(\lambda_Q + \lambda_P)} \frac{\lambda_Q^q}{q!} \frac{\lambda_P^p}{p!} \\ &= e^{-(\lambda_Q + \lambda_P)} \sum_{q=0}^{\infty} \sum_{p=0}^{\infty} \frac{\lambda_Q^q}{q!} \frac{\lambda_P^p}{p!} \left( [f(q+1, p) - f(q, p)] r q_0 + [f(q-1, p) - f(q, p)] r q \right. \\ &\quad \left. + [f(q-1, p+1) - f(q, p)] k_2 q + [f(q+1, p-1) - f(q, p)] k_1 p \right). \end{aligned}$$

In the two sums the factor multiplying  $f(q, p)$  for any  $(q, p) \in \mathbb{N} \times \mathbb{N}$  comes only from terms involving  $\{q - 1, q, q + 1\}$  and  $\{p - 1, p, p + 1\}$  and equals  $e^{-(\lambda_Q + \lambda_P)}$  times

$$\begin{aligned} & \frac{\lambda_Q^{q-1}}{(q-1)!} \frac{\lambda_P^p}{p!} r q_0 - \frac{\lambda_Q^q}{q!} \frac{\lambda_P^p}{p!} r q_0 + \frac{\lambda_Q^{q+1}}{(q+1)!} \frac{\lambda_P^p}{p!} r(q+1) - \frac{\lambda_Q^q}{q!} \frac{\lambda_P^p}{p!} r q \\ & + \frac{\lambda_Q^{q+1}}{(q+1)!} \frac{\lambda_P^{p-1}}{(p-1)!} k_2(q+1) - \frac{\lambda_Q^q}{q!} \frac{\lambda_P^p}{p!} k_2 q \\ & + \frac{\lambda_Q^{q-1}}{(q-1)!} \frac{\lambda_P^{p+1}}{(p+1)!} k_1(p+1) - \frac{\lambda_Q^q}{q!} \frac{\lambda_P^p}{p!} k_1 p \\ & = \frac{\lambda_Q^q}{q!} \frac{\lambda_P^p}{p!} \left( \frac{q}{\lambda_Q} r q_0 - r q_0 + \frac{\lambda_Q}{q+1} r(q+1) - r q + \frac{\lambda_Q}{q+1} \frac{p}{\lambda_P} k_2(q+1) \right. \\ & \quad \left. - k_2 q + \frac{q}{\lambda_Q} \frac{\lambda_P}{p+1} k_1(p+1) - k_1 p \right) \\ & = \frac{\lambda_Q^q}{q!} \frac{\lambda_P^p}{p!} \left( \frac{q}{q_0} r q_0 - r q_0 + q_0 r - r q \right) = 0 \end{aligned}$$

since  $\lambda_Q = q_0$  and  $\lambda_P/\lambda_Q = k_2/k_1$ .

Thus, in equilibrium the input rate for  $(Q_2, P_2)$ , which is  $rQ_1$ , has a Poisson distribution with mean  $r q_0$  and is independent of  $P_1$ . Let  $\pi_{q_0}(q_1) \otimes \pi_{\lambda_Q}(q_2) \otimes \pi_{\lambda_P}(p_2)$  be a product of three Poisson distributions with rates  $q_0, \lambda_Q = q_0$ , and  $\lambda_P = q_0 k_2/k_1$ , respectively. To show that this is a stationary distribution for the process  $(Q_1, Q_2, P_2)$  we need to check that

$$\sum_{q_1=0}^{\infty} \sum_{q_2=0}^{\infty} \sum_{p_2=0}^{\infty} \mathcal{A}_{q_1} f(q_2, p_2) \pi_{q_0}(q_1) \pi_{q_1}(q_2, p_2) = 0$$

for any choice of function  $f \in \mathcal{D}(\mathcal{A}_{q_1})$ . For each fixed value  $q_1$ , according to our previous calculation the inner two sums give 0, so the whole sum is 0.

Thus, in equilibrium,  $(Q_2, P_2)$  have the distribution  $\pi_{\lambda_Q}(q) \otimes \pi_{\lambda_P}(p)$  with  $\lambda_Q = q_0, \lambda_P = q_0 k_2/k_1$  as well, and are independent of  $(Q_1, P_1)$ . It follows by induction that the stationary distributions for  $\{(Q_i, P_i)\}$  are i.i.d. as  $\pi_{\lambda_Q}(q) \otimes \pi_{\lambda_P}(p)$ , with  $\lambda_Q = q_0, \lambda_P = q_0 k_2/k_1$ . This is also an example of a clustering process satisfying the detailed balance conditions with linear rates discussed in section 8.2 of [24].  $\square$

We point out that the mean number of particles both on- and off-transport is  $(1 + \frac{k_2}{k_1})q_0 = (1 + \frac{k_2}{k_1})\rho_0\delta$ , scaling with the size of a compartment. To obtain the mean number of particles per unit length we add particles in  $\approx 1/\delta$  compartments, and the mean number of particles per unit length is  $(1 + \frac{k_2}{k_1})\rho_0$  independent of the choice of compartment size.

One immediate consequence is the analogous result for the number of particles in the stochastic noncompartmental model at any location along the axon. Namely, suppose the particles move according to the piecewise-linear Markov process  $(X, \xi)$  with a Poisson rate  $\rho_0$  influx of new particles at location 0. Then, at any location  $0 < x < L$  along the axon, the numbers of particles  $(Q_{(x, x+dx)}, P_{(x, x+dx)})$  on- and off-transport, respectively, have the stationary distribution  $\text{Pois}(\rho_0 dx) \otimes \text{Pois}(\frac{k_2}{k_1} \rho_0 dx)$ , where for any  $x_1, \dots, x_k \in (0, L)$ ,  $\{Q_{x_i}\}_{1 \leq i \leq k}$  and  $\{P_{x_i}\}_{1 \leq i \leq k}$  are mutually independent. We note that this result would not have been obvious without going through the compartmental model first, yet its consequences for prediction and analysis of the long-term stochastic behavior of the system are quite powerful.

**4.2. Homogeneity of the axons at equilibrium.** Recall that  $\delta = 10nm$ , roughly the step size of motor proteins, and that axons can be up to one meter in length. Thus we are interested in phenomena on all the length scales  $10^\nu\delta$ , where  $\nu = 0, 1, 2, \dots, 8$ . Let  $\Delta = 10^\nu\delta$ ; we want to determine how similar different segments of the axon of size  $\Delta$  are. Let  $Q_\Delta$  and  $P_\Delta$  denote the numbers of on-track and off-track particles in a segment of length  $\Delta$ .

In equilibrium,  $Q_\Delta$  and  $P_\Delta$  are both sums of  $10^\nu$  independent Poisson random variables with parameters  $\lambda_Q = q_0$  and  $\lambda_P = \frac{k_2}{k_1}q_0$ , respectively. Therefore, the distributions of  $Q_\Delta$  and  $P_\Delta$  are Poisson with parameters  $10^\nu\lambda_Q$  and  $10^\nu\lambda_P$ , respectively. The mean and the variance of the number of particles in the segment of length  $\Delta$  are  $10^\nu(\lambda_Q + \lambda_P)$ . To see how homogeneous different slices of length  $\Delta$  are, we consider the coefficient of variation,  $c_\Delta$ , which is the standard deviation divided by the mean:

$$c_\Delta = \frac{1}{\sqrt{(\lambda_Q + \lambda_P)10^\nu}} = \frac{1}{\sqrt{(1 + k_2/k_1)q_010^\nu}}.$$

As indicated in section 2,  $q_0$  is in the range 0.25 to 2.5 in different axons. For illustrative purposes here, we will assume  $q_0 = 1$ . Since  $k_2/k_1 = 3$ , we see that the scale-dependent coefficient of variation  $c_\Delta = 1/(2\sqrt{10^8\Delta})$ . Therefore, at the ten nanometer scale the coefficient of variation is simply 1/2. At the micron scale  $c_\Delta = 1/20$ , and at the millimeter scale  $c_\Delta = 0.5 \times 10^{-5/2}$ . The cutoff between “high variance” and “low variance” distributions is usually considered to be when the coefficient of variation is near 1, so by this standard the axon is extremely homogeneous in its length at large scales.

**4.3. Balance between efficient transport and targeted delivery.** The preceding characterization of the transport apparatus enables us to address questions concerning whole cell function. One core issue is that intracellular transport must simultaneously accommodate two functional demands: some material, such as the enzymes used to construct neurotransmitters, must be transported from the soma to the axon terminal in a timely manner; whereas other cargo, such as sodium channels, need to be delivered to unspecified locations as needs arise throughout the length of the cell. The tradeoff between these two goals is clear. If a typical vesicle spends the vast majority of its time in transport mode, the mean velocity will be close to the mean on-transport velocity, but any needs that arise in the central part of the cell will be neglected. On the other hand, if a typical vesicle spends too much time off transport, presumably available for use if needed locally, then it will take substantially longer to traverse the entire cell.

Recently Bressloff and Newby [30], [31] modeled particles that are created near the nucleus that then undergo intermittent search (being in search mode while off-transport and not in search mode when on-transport) for a target hidden somewhere along the axon. Their model is a noncompartmental individual particle model with the additional feature that vesicles can move backward as well as forward. They compute the probability that the particle is successful and conditioning on success, the mean first hitting time. With our systemwide model we can accommodate the observation that if a given vesicle misses the target, another vesicle with similar cargo will pass by before too long. We will assess this hitting time under two assumptions about the density of relevant material. Our standing assumption  $q_0 \sim 1$  is appropriate for types of cargo that are found densely throughout the cell. In this setting, the wait time is essentially just the time it takes for one of several nearby vesicles to unbind from transport in the target region. A more interesting case is a setting where the



needed cargo is sparsely distributed, say,  $q_0 \sim 10^{-3}$ . In this setting, if the first cargo to reach the target region fails to unbind, there will be significant time before the next arrival. As we will see, we can still assess the trade-off intrinsic between risking a target miss and diminishing the time of the next arrival.

To make the discussion concrete we define a target region  $R_n = \{i_* + 1, \dots, i_* + n\}$ , where  $i_* \in 1, \dots, N - n$ . At time zero, we take the system  $\{Q_i(0), P_i(0)\}$  to be drawn from the stationary distribution described by Proposition 4.1 conditioned on the event that for all  $i \in R_n$ ,  $Q_i(0) = P_i(0) = 0$ . We introduce the hitting time  $H_n := \inf\{t > 0 : \sum_{i \in R_n} P_i(t) > 0\}$ , which marks the first time a particle is off-transport while in  $R_n$ . Since computing the mean of  $H_n$  is analytically intractable, we introduce another hitting time  $H'_n$ , stochastically dominating  $H_n$ , that nevertheless reflects the essential trade-off between maximizing mean velocity and making detachment from transport likely in the target region.

Let  $I_* := \{i \in \{1, \dots, i_*\} \mid Q_i(0) + P_i(0) > 0\}$  be the set of all nonempty sections of the cell at time 0. Among the particles in these sections some will be “successful” in that they will detach from transport in the target region, while others will be unsuccessful. Labeling each particle in these terms, we decompose  $I_0$  into locations with successful particles  $I_*^s$  and unsuccessful particles  $I_*^u$ . We are interested in the time  $H'_n$  at which the rightmost successful particle, that is, a particle starting from position  $i_m = \max\{I_*^s\}$ , detaches while in  $R_n$ . If  $S = 0$ , then we define  $i_m = \max\{\emptyset\} := 0$ . For the position of this particle we use the notation  $(X_t^\delta)_{t \geq 0}$  from section 3, where  $X_0^\delta = \delta i_m$ ,  $X_t^\delta \in \{\delta i_m, \dots, \delta N\}$ , together with the indicator  $(\xi_t^\delta)_{t \geq 0}$ ,  $\xi_t^\delta \in \{0, 1\}$ , of whether the particle is on- or off-transport, respectively. Let

$$H'_n := \inf\{t > 0 : (X_t^\delta, \xi_t^\delta) = (\delta(i_* + 1), 1)\}.$$

Note that  $H'_n = \inf\{t > 0 : X_t^\delta \in \delta R_n\}$ , since particles cannot skip sections and always enter a section on transport. If we were to analogously define a sequence of times  $\{H_n^i\}_{i \in I_*^s}$ , where for each  $i \in I_*^s$ ,  $H_n^i = \inf\{t > 0 : X_t^\delta \in \delta R_n\}$  with  $X_0^\delta = \delta i$  and  $\xi_0^\delta \in \{0, 1\}$ , then the exact first hitting time of the target region will satisfy  $H_n = \min\{H_n^i : i \in I_*^s\}$ . Hence, clearly  $H_n \leq H_n^{i_m} \equiv H'_n$ .

Since the exact distribution of  $H_n^i$  is complicated,  $\mathbb{E}[H_n^i]$  is analytically intractable, and instead we focus on finding a simple expression for  $\mathbb{E}[H'_n]$ . Our point is that, in the sparse material limit ( $q_0$  small), the rightmost particle becomes increasingly likely to be the first successful particle to detach in the target region, and  $H_n$  approaches  $H'_n$ .

The computation of  $\mathbb{E}[H'_n]$  requires computing the time it takes a particle to travel a certain axonal distance, given by the following.

LEMMA 4.2. *Let the initial position of the particle be  $(X_0^\delta, \xi_0^\delta) = (0, 1)$ , and let  $L_* \in \{1, \dots, L\}$  be given, where  $L$  is the total length of the axon. Then the time  $T_{L_*} = \inf\{t > 0 : X_t^\delta = L_*\}$  satisfies*

$$\mathbb{E}[T_{L_*}] = \frac{L_*(k_2 + k_1)}{vk_1}.$$

*Proof.* Since  $(X_t^\delta, \xi_t^\delta)_{t \geq 0}$  is a Markov process with generator

$$\begin{aligned} A^\delta f(x, \xi) &= r\xi[f(x + \delta, \xi) - f(x, \xi)] \\ &\quad + k_2\xi[f(x, \xi - 1) - f(x, \xi)] + k_1(1 - \xi)[f(x, \xi + 1) - f(x, \xi)], \end{aligned}$$

it follows that  $M_t^1 := X_t^\delta - r\delta \int_0^t \xi_s^\delta ds$  and  $M_t^2 := \xi_t^\delta + \int_0^t k_2 \xi_s^\delta ds - \int_0^t k_1(1 - \xi_s^\delta) ds$  are both martingales. Since both  $M^1$  and  $M^2$  have bounded increments and  $\mathbb{E}[T_{L_*}] < \infty$ ,

the optional stopping theorem implies that

$$0 = \mathbb{E}[M_0^1] = \mathbb{E}[M_{T_{L_*}}^1] = L_* - v\mathbb{E}\left[\int_0^{T_{L_*}} \xi_s^\delta ds\right] \Rightarrow E\left[\int_0^{T_{L_*}} \xi_s^\delta ds\right] = L_*/v$$

and

$$1 = \mathbb{E}[M_0^2] = \mathbb{E}[M_{T_{L_*}}^2] = 1 - k_1\mathbb{E}[T_{L_*}] + (k_2 + k_1)\mathbb{E}\left[\int_0^{T_{L_*}} \xi_s^\delta ds\right],$$

which implies

$$\mathbb{E}[T_{L_*}] = \frac{k_2 + k_1}{k_1} \mathbb{E}\left[\int_0^{T_{L_*}} \xi_s^\delta ds\right],$$

and our claim follows.  $\square$

We also note that the same computation holds for the mean time a particle in the stochastic noncompartmental model  $(X_t, \xi_t)_{t \geq 0}$  takes to reach a distance  $L_*$ , as the two martingales used in the proof depend only on  $v$  and not on  $\delta$ .

We next compute the hitting time  $H'_n$  of a particle started at location  $i_m \in I_*$ .

LEMMA 4.3. *Let the system  $\{(Q_i(0), P_i(0)), i = 1, \dots, N\}$  have the stationary distribution given by Proposition 4.1 conditional on  $Q_i(0) = P_i(0) = 0$  for all  $i \in R_n$ . Then*

$$\mathbb{E}[H'_n] = (1 - e^{-\lambda_n}) \left[ \frac{1}{r q_0 p_n} + \frac{k_2}{k_1(k_2 + k_1)} \right] + \frac{1}{k_2 p_n} \left[ 1 - \left( \frac{r}{k_2 + r} \right)^n \left( 1 + \frac{n k_2}{k_2 + r} \right) \right],$$

where  $p_n = 1 - \left(\frac{r}{k_2+r}\right)^n$  and  $\lambda_n = \frac{k_2+k_1}{k_1} q_0 i_* p_n$ .

*Proof.* The proof of Proposition 4.1 shows that conditioning on the values of  $P_i(0)$ ,  $i \in R^n$ , does not affect the law of  $\{(Q_i(t), P_i(t)), i \leq i_*\}$ ; hence for any  $t \geq 0$ , they form two mutually independent sequences of  $\text{Pois}(q_0)$  and  $\text{Pois}\left(\frac{k_2 q_0}{k_1}\right)$  random variables.

Let  $S$  denote the number of successful particles between sites 0 and  $i_*$  at time zero. The total of number of particles at time 0 at these sites that are either on- or off-transport is distributed as a  $\text{Pois}\left(\frac{k_2+k_1}{k_1} q_0 i_*\right)$  variable. We next compute the probability  $p_n$  that any particle once it reached the target region is “successful” in detaching there. This probability can be written  $p_n := \sum_{i=1}^n p(i)$ , where  $p(i)$  is the probability it first detaches at location  $i_* + i$ . Since  $p(i) = \left(\frac{r}{k_2+r}\right)^{i-1} \frac{k_2}{k_2+r}$  is the chance a particle gets off-transport in the  $i$ th compartment,  $p_n = \sum_{i=1}^n p(i) = 1 - \left(\frac{r}{k_2+r}\right)^n$ . The probability that a particle is successful does not depend on its location at time 0 or whether it was on- or off-transport at that time. Hence,  $S$  is distributed as a  $\text{Pois}\left(\frac{k_2+k_1}{k_1} q_0 i_* p_n\right)$  variable, and conditioned on the value  $S$ , the set of locations of the particles at time 0 is distributed as a set of  $S$  draws from the uniform distribution on  $\{1, \dots, i_*\}$ . Since  $i_m$  is the maximum of  $S$  samples from a  $\text{Uniform}\{1, \dots, i_*\}$  distribution, we have  $\mathbb{E}[i_m | S] = i_* - \frac{1}{i_*^S} \sum_{x=1}^{i_*-1} x^S \approx i_* \frac{S}{S+1}$ .

We now decompose  $H'_n = H'_{e,n} + H'_{o,n}$ , where we let  $H'_{e,n}$  be the time it takes a successful particle initially at location  $i_m$  to enter the region  $R_n$ , and  $H_{o,n}$  is the time it takes any successful particle after it enters the region to get off-transport. Thus,  $\mathbb{E}[H'_n] = \mathbb{E}[H'_{e,n}] + \mathbb{E}[H'_{o,n}]$ .

If a successful particle starts at location  $X_0^\delta = i_m \in \{0, \dots, i_*\}$  on transport  $\xi_0^\delta = 1$ , then the time it takes to enter the region  $R^n$  is by Lemma 4.2

$$\mathbb{E}[H'_{e,n} | (X_0^\delta, \xi_0^\delta) = (i_m, 1)] = \frac{(i_* - i_m)\delta(k_2 + k_1)}{vk_1}.$$

If it starts at location  $X_0^\delta = i_m \in \{1, \dots, i_*\}$  off transport  $\xi_0^\delta = 1$ , then it takes an additional exponential time with mean  $1/k_1$  for it to get back on transport at the same location, so  $\mathbb{E}[H'_{e,n} | (X_0^\delta, \xi_0^\delta) = (i_m, 0)] = \mathbb{E}[H'_{e,n} | (X_0^\delta, \xi_0^\delta) = (i_m, 1)] + \frac{1}{k_1}$ , and since we assume new particles always enter the system on-transport,

$$\mathbb{E}[H'_{e,n} | i_m] = \frac{(i_* - i_m)\delta(k_2 + k_1)}{vk_1} + \frac{1}{k_1} \mathbf{1}_{i_m > 0}.$$

Since  $\mathbb{E}[i_m | S] = i_* \frac{S}{S+1}$ ,  $\mathbb{P}\{i_m > 0\} = \mathbb{P}\{S > 0\}$ , we get that

$$\mathbb{E}[H'_{e,n}] = \mathbb{E}\left[\frac{1}{S+1}\right] \frac{i_*(k_2 + k_1)}{rk_1} + \frac{k_2}{k_1(k_2 + k_1)} \mathbb{P}\{S > 0\}$$

with  $S \sim \text{Pois}(\lambda_n)$ ,  $\lambda_n = \frac{k_2+k_1}{k_1} q_0 i_* p_n$ . Since  $\mathbb{E}[S] = \frac{1-e^{-\lambda_n}}{\lambda_n}$ ,  $\mathbb{P}\{S > 0\} = 1 - e^{-\lambda_n}$ , we get

$$\mathbb{E}[H'_{e,n}] = (1 - e^{-\lambda_n}) \left( \frac{1}{rq_0 p_n} + \frac{k_2}{k_1(k_2 + k_1)} \right).$$

To calculate  $\mathbb{E}[H'_{o,n}]$  note that if a particle first gets off-transport in the  $i$ th compartment, then the time of its travel until this point is a sum of  $i$  i.i.d. exponential random variables with parameter  $k_2 + r$ . The probability a successful particle first gets off-transport in the  $i$ th compartment is  $\frac{p(i)}{p_n}$ . Hence, the time a successful particle takes to get off-transport once it enters the region  $R$  has the mean

$$\mathbb{E}[H'_{o,n}] = \sum_{i=1}^n \frac{p(i)}{p_n} \frac{i}{k_2 + r} = \frac{1}{k_2 p_n} \left[ 1 - \left( \frac{r}{k_2 + r} \right)^n \left( 1 + \frac{nk_2}{k_2 + r} \right) \right],$$

and our claim follows.  $\square$

To complete our analysis, we wish to characterize this result in terms of length along the axon and independent of the stepping size  $\delta$ . To this end, we fix a length  $\ell$ , and for a given compartment size  $\ell$ , we let  $n = \lceil \ell/\delta \rceil$ , and for the start of the region we let  $i_* = \lceil \ell_*/\delta \rceil$ . As such, as  $\delta \rightarrow 0$  the limiting region becomes  $R_\ell = (\ell_*, \ell_* + \ell)$ . Then, under the assumption that  $q_0/\delta \rightarrow \rho_0$  we have

$$p_n \rightarrow p_\ell = 1 - e^{-\ell k_2/v}, \quad \lambda_n \rightarrow \lambda_\ell = \frac{k_2 + k_1}{k_1} \rho_0 \ell_* p_\ell,$$

and  $\mathbb{E}[H'_n] \rightarrow \mathbb{E}[H']$ , where

$$(4.1) \quad \mathbb{E}[H'] = (1 - e^{-\lambda_\ell}) \left[ \frac{1}{v\rho_0 p_\ell} + \frac{k_2}{k_1(k_2 + k_1)} \right] + \frac{1}{k_2 p_\ell} \left[ 1 - e^{-\frac{\ell k_2}{v}} \left( 1 + \frac{\ell k_2}{v} \right) \right].$$

It remains to interpret this result with respect to the parameter choices we have made. First, we must set a value for the size  $\ell$  of the target region. For this purpose we note that the typical size of a node of Ranvier—a gap in the myelin sheath of a

myelinated axon where sodium channels are concentrated in the cell membrane—is approximately one micron.<sup>1</sup>

As seen in the preceding proof there are three contributions to  $\mathbb{E}[H']$ , and we next analyze them in terms of their dominance for the overall value. We begin with the contribution from the last term, which is the time it takes for a successful particle to detach once it has reached the target region. By our choice of  $\ell$ , the recurring ratio  $\ell/v$  is one. Along with the earlier assumption that  $k_2 = 1$ , the entire term simplifies to  $(e-2)/(e-1) = 0.4$  s. The multiplicative factor  $1 - e^{-\lambda_\ell}$  preceding the first term of (4.1) results from a boundary effect: when  $\ell_*$  is very close to zero, it is very unlikely there are any particles already in the system between the soma and the target region. When the target region is in the middle of the axon, this contribution from particles in that section of the axon at time 0 is significant.

When  $q_0 \sim 1$  as assumed earlier, then  $\rho_0 = q_0/\delta \sim 10^8$ . Then if  $\ell_* > 10^{-8}$ , which corresponds to the target region being just one motor step down the length of the axon, we have  $\rho_0 \ell_* \sim 1$  and  $\lambda_\ell > 2.3$ , so  $1 - e^{-\lambda_\ell} > 0.9$ . Looking at the first term inside the parenthesis, we note that  $v\rho_0 \sim 10^2$ , while  $p_\ell = 1 - \exp(-\ell k_2/v) \approx 0.6$  is the probability that any given vesicle will be successful in detaching from transport in the target region, so  $1/(v\rho_0 p_\ell) \sim 10^{-2}$ . Meanwhile, the term  $k_2/(k_1(k_2 + k_1))$ , which is the expected time to bind to transport if a successful particle happens to be off transport at time 0, is  $1/4$  for the assumed values of  $k_2$  and  $k_1$ . Therefore, under the  $q_0 \sim 1$  assumption, both this and the final term contribute significantly to the hitting time.

However, in the sparse material regime, say,  $q_0 \sim 10^{-3}$ , we have then  $\rho_0 \sim 10^5$ , and the factor  $1 - e^{-\lambda_\ell} > 0.9$  when  $\ell_* > 10^{-5}$ , that is, if the target region is at least  $1/100$ , or at least 1000 segments, down the length of a  $10^{-3}m$  axon. However, now the rate limiting factor is the wait time for the first successful vesicle to arrive in the target region, which is captured by the term  $1/v\rho_0 p$ . The product  $v\rho_0 \approx 0.1$  measures the average rate at which new particles should arrive, and together with the probability of success  $p_\ell \approx 0.6$ , we now have  $1/v\rho_0 p_\ell \approx 16$  s. So, in the sparse material regime the first summand in  $\mathbb{E}[H']$  giving the mean time for arrival of the particles to the target region dominates.

It is interesting to consider what happens to the arrival rate term under perturbations of the various parameters. In particular, we draw the reader's attention to changes in  $v$ . When viewing intermittent search in terms of a single particle, the probability of finding the target is strictly decreasing in  $v$ . Higher velocity seems to be the enemy of finding the target. Indeed, from a system point of view, this implies that the system will require more trials before a successful particle arrives in the target region. What (4.1) gives us is the ability to assess how much more quickly the trials will happen. In fact, the function  $v(1 - \exp(-\ell k_2/v))$  is increasing in  $v$ . Therefore, the entire expected wait time  $\mathbb{E}[H']$  is actually *decreasing* in  $v$ . That is to say, while the particles are less likely to succeed, they will be arriving rapidly enough to counterbalance the lost time. We believe that this kind of quantitative analysis will prove fruitful in future study when coupled with more details of the biology of deposition of materials in the cell membrane.

---

<sup>1</sup>We do not wish to claim that this is a complete model for deposition of sodium channels in nodes of Ranvier, as the particulars of the biology—which may include factors like local signals that encourage motors to detach from microtubules near the nodes—are not fully understood. We merely wish to use the size of the nodes to fix our intuition about the size of a target to other important length scales, such as that of microtubules, which are also a micron in length.

**4.4. Approaching equilibrium.** We have seen above that the axon is very homogeneous at stochastic equilibrium on a space scale down to micrometers. One of the beautiful properties of transport with reversible binding is that if it is locally out of equilibrium, the on-off dynamics can return the system to equilibrium on a much faster time scale than waiting for new material to arrive from the nucleus. Furthermore, as discussed in [44], one of the key goals of biophysics investigation is the discovery of behaviors for which biochemical regulation is not necessary. This is of fundamental importance to the biological function of the system because it means that the axon will automatically “repair” itself without central control of the repair process.

How good is this mechanism? If a segment of the axon is far away from equilibrium, how long does it take to get back close to equilibrium? To investigate this question, we imagine that the axon is at stochastic equilibrium except for some segment  $R$ , where the total number of particles on- and off-transport,  $Q_i(0) + P_i(0) = 0$ , for all  $i \in R$ , is zero at time 0. Let  $a$  and  $\varepsilon$  be given small numbers and suppose that  $\lambda_\infty$  is the mean vector for  $(Q_i, P_i)$  at stochastic equilibrium. We want to compute (an upper bound for) the time  $t^*$  so that

$$\mathbb{P}\{|(Q_i(t^*), P_i(t^*)) - \lambda_\infty| \geq a|\lambda_\infty| \forall i \in R\} \leq \varepsilon,$$

that is, the probability that  $(Q_i, P_i)$  for all  $i \in R$  is significantly different from  $\lambda_\infty$  is very small. In the applications below we will choose  $a = 0.1$  and  $\varepsilon = 0.05$ , and we will see that a 10 micron segment can recover in about 10 seconds, while a 1 millimeter segment will take about 1000 seconds or 15 minutes to recover. We study this question first for a single location  $R = \{i\}$  and then use the estimates derived to scale the results to segments of any length.

**PROPOSITION 4.4.** *Let  $(Q_i(0), P_i(0)) = (0, 0)$ , and let the constants  $a > 0$  and  $\varepsilon \in (0, 1)$  satisfy the relationship  $a^2\varepsilon|\lambda_\infty| > 1$ , where  $\lambda_\infty = q_0(1, \frac{k_2}{k_1})$  is the equilibrium vector of  $(Q_i, P_i)$ . Then there exists  $t_* > 0$  such that for all  $t \geq t_*$ ,*

$$\mathbb{P}\{|(Q_i(t), P_i(t)) - \lambda_\infty| \geq a|\lambda_\infty|\} \leq \varepsilon.$$

*In fact, the choice  $t_* = \alpha^{-1} \ln(\frac{\sqrt{p|\lambda_\infty|}}{a\sqrt{\varepsilon|\lambda_\infty|-1}})$  is sufficient, where*

$$\alpha := \frac{1}{2} \left( k_2 + k_1 + r - \sqrt{(k_2 + k_1 + r)^2 - 4k_1r} \right).$$

We note that given a particular choice of  $a$  and  $\varepsilon$ , the condition  $a^2\varepsilon|\lambda_\infty| > 1$  guarantees that there are “enough” particles.

*Proof.* Note that for any given  $\beta \in (0, 1)$ , we may choose  $t_* > 0$  such that for all  $t \geq t_*$ , the vector of means  $\lambda(t) := \mathbb{E}[(Q_i(t), P_i(t))]$  satisfies  $|\lambda(t) - \lambda_\infty| \leq a\beta|\lambda_\infty|$ . Then

$$\begin{aligned} \mathbb{P}\{|(Q_i(t), P_i(t)) - \lambda_\infty| \geq a|\lambda_\infty|\} &\leq \mathbb{P}\{|(Q_i(t), P_i(t)) - \lambda(t)| + |\lambda(t) - \lambda_\infty| \geq a|\lambda_\infty|\} \\ &\leq \mathbb{P}\{|(Q_i(t), P_i(t)) - \lambda(t)| \geq a(1 - \beta)|\lambda_\infty|\}. \end{aligned}$$

Applying Chebyshev’s inequality and observing that the variance of a Poisson random variable is equal to its mean, we conclude that

$$\mathbb{P}\{|(Q_i(t), P_i(t)) - \lambda_\infty| \geq a|\lambda_\infty|\} \leq \frac{\text{Var}[(Q_i(t), P_i(t))]}{a^2(1 - \beta)^2|\lambda_\infty|^2} = \frac{|\lambda(t)|}{a^2(1 - \beta)^2|\lambda_\infty|^2}.$$

Since the initial conditions for both  $P_i$  and  $Q_i$  are less than their respective equilibrium values, each are monotonically increasing functions, and the above reduces to

$$\mathbb{P}\{|(Q_i(t), P_i(t)) - \lambda_\infty| \geq a|\lambda_\infty|\} \leq \frac{1}{a^2(1 - \beta)^2|\lambda_\infty|}$$

for all  $t > t_*$ . In order to satisfy the requirement that the right-hand side must be less than  $\varepsilon$ , we solve for  $\beta$  and find  $\beta = 1 - (a\sqrt{\varepsilon|\lambda_\infty|})^{-1}$  provided that  $a\sqrt{\varepsilon|\lambda_\infty|} > 1$ .

It remains to study the convergence of the mean and the appropriate choice of  $t_*$ . The dynamics of the mean vector  $\lambda(t)$  are given by the ODE

$$(4.2) \quad \frac{d}{dt}\lambda(t) = -A_1\lambda(t) + q_0r\mathbf{e}_1,$$

where  $\mathbf{e}_1$  is the unit vector  $(1, 0)$  and  $A_1 = \begin{pmatrix} k_2+r & -k_1 \\ -k_2 & k_1 \end{pmatrix}$ .

The solution to (4.2) is  $\lambda(t) = \lambda_\infty + e^{-A_1t}(\lambda(0) - \lambda_\infty)$ , where  $\lambda_\infty = q_0rA_1^{-1}\mathbf{e}_1 = q_0(1, \frac{k_2}{k_1})$ . This yields the estimate

$$|\lambda(t) - \lambda_\infty| \leq |e^{-A_1t}\lambda_\infty| \leq e^{-\alpha t}|\lambda_\infty|,$$

where  $\alpha$  is the smaller of the eigenvalues of  $A_1$ . Noting that  $\alpha > 0$ ,  $t_*$  may be chosen so that  $e^{-\alpha t_*} \leq a\beta$ ; i.e.,  $t_* = \alpha^{-1} \ln(1/(a\beta))$  is an upper bound for the time to be close to equilibrium with high probability.  $\square$

In order to calculate return to equilibrium at various scales, we now suppose that the whole axon is in statistical equilibrium except for a segment  $R$  of length  $\Delta = \delta 10^\nu$ , in which we will assume that there are no particles either on or off the track. Proposition 4.4 covered the case  $\nu = 0$ . We are interested in  $\nu = 1, \dots, 8$ . We imagine that the axon is broken up into  $10^{8-\nu}$  segments of length  $\Delta$ . In this rescaled system, the unbinding and binding rates per particle,  $k_2$  and  $k_1$ , remain the same, as well as the mean on-transport velocity  $v$ . In order to retain this mean velocity, the rate of lateral stepping must be decreased to  $\tilde{r} = r10^{-\nu}$ .

The ODE for the mean vector of the rescaled system is given by

$$(4.3) \quad \frac{d}{dt}\tilde{\lambda}(t) = -\tilde{A}_1\tilde{\lambda}(t) + q_0r\mathbf{e}_1$$

with  $\tilde{A}_1 = \begin{pmatrix} k_2+\tilde{r} & -k_1 \\ -k_2 & k_1 \end{pmatrix}$ .

We note that the last term in (4.3) contains an  $r$  rather than an  $\tilde{r}$ . This is because the input rate is unchanged while the exit rate is diminished. The resulting equilibrium value is therefore rescaled as well:  $\tilde{\lambda}_\infty = q_0r\tilde{A}_1^{-1}\mathbf{e}_1 = q_0\frac{r}{\tilde{r}}\begin{pmatrix} 1 \\ k_2/k_1 \end{pmatrix} = q_010^\nu\begin{pmatrix} 1 \\ k_2/k_1 \end{pmatrix}$ . Both components of this vector are of order  $10^\nu$ , as expected.

Using the parameters discussed in section 2,  $k_2 = 1$ ,  $k_1 = \frac{1}{3}$ ,  $v = 10^{-6}m/s$ ,  $r = 10^2s^{-1}$ . This implies  $\tilde{r} = 10^{2-\nu}s^{-1}$ . We choose the thresholds to be  $a = 0.1$  and  $\varepsilon = 0.05$ , so the constraint that  $a^2\varepsilon|\lambda_\infty| > 1$  requires that  $|\lambda_\infty| > 2 \times 10^3$ . Since  $0.25 \leq q_0 \leq 2.5$ , this is indeed the case if  $\nu > 3$ , i.e., if the segment has length greater than 10 microns.

It follows from Proposition 4.4 that the time to equilibrium,  $\tilde{t}_*$ , is proportional to  $\tilde{\alpha}^{-1}$ , where  $\tilde{\alpha}$  satisfies

$$\begin{aligned} \tilde{\alpha} &= \frac{1}{2} \left( k_2 + k_1 + \tilde{r} - \sqrt{(k_2 + k_1 + \tilde{r})^2 - 4k_1\tilde{r}} \right) \\ &= \frac{2k_1\tilde{r}}{k_2 + k_1 + \tilde{r} + \sqrt{(k_2 + k_1 + \tilde{r})^2 - 4k_1\tilde{r}}}. \end{aligned}$$

For the given parameter values (with  $|\lambda_\infty| > 5 \times 10^3$ , in particular), the constant of proportionality  $\ln(\sqrt{\varepsilon|\lambda_\infty|}/(a\sqrt{\varepsilon|\lambda_\infty|} - 1))$  is contained in the interval (2.3, 2.5) and does not have an impact on how the relaxation time scales with  $\nu$ . In terms of analyzing  $\tilde{\alpha}$ , we note that  $k_1\tilde{r}$  is small compared to  $k_2$ . To leading order,  $\tilde{\alpha} \sim (k_1\tilde{r})/(k_2 + k_1 + \tilde{r}) \sim 10^{2-\nu}s^{-1}$ . It follows that  $\tilde{t}_* \sim 10^{\nu-2}s$ . Thus, for a 10 micron segment ( $\nu = 3$ ) the time to recover is about 10 seconds, and for a 1 millimeter segment ( $\nu = 5$ ) the time to recover is 1000 seconds or 15 minutes. The time to recover depends, of course, on the parameter  $\varepsilon$  that represents what we mean by “close.” We also note that one can compute various measures of time to recover using the PDE models discussed in section 3.

**5. Discussion.** In this paper, we created a spatial Markov-chain model for studying various aspects of fast axonal transport. Previous models that use PDEs treat the velocity of transport as constant when particles are attached to the fast transport system. Since it is known that transport along the microtubules is itself stochastic, it is important to have a fully stochastic model. Our model allows us to unify and extend previous work. In section 3.2 we show that from the particle perspective as the compartment size  $\delta$  tends to zero, our model converges in distribution on the Skorokhod space of cadlag functions to the piecewise-deterministic model analyzed by Brooks [5]. Namely, we show that the paths of particles in our model converge to those of particles in a stochastic noncompartmental model. The argument proceeds by an explicit computation of the finite dimensional distributions and a tightness argument. In Proposition 3.4 we give a rigorous probabilistic proof of why the paths of particles follow “approximate traveling waves” described by other authors [35], [5], [13]. This proof is based on stochastic averaging arguments which show that a functional central limit theorem holds on the space of continuous functions for the paths of particles as the compartment size decreases. The diffusion of particles around their mean position can consequently be approximately described jointly for all time by a Brownian motion with the appropriate diffusion coefficient.

In section 2, we show how to use existing experimental data to identify (ranges for) all the parameters of our model. In light of this, we can use the model to investigate several important biological questions. These are based on describing the spatial distribution of multiple particles in our model. In section 4.1 we derive the stationary distribution for the number of particles in different compartments on- and off-transport along the axon. This gives an explicit description of the stochasticity of the system that is present even after a long time. In section 4.2 we derive estimates for how homogeneous the axon is on different spatial scales. In section 4.3 we study a question introduced by Bressloff by providing a stochastic quantity which describes the balance the system needs to achieve between rapid transport that brings new material quickly and efficient local search that improves time of delivery to a target. Finally, in section 4.4, we use the model to calculate the length of time that it would take for axonal segments of different lengths to recover to near stochastic equilibrium after they have been depleted of vesicles.

In our stochastic compartmental model all event wait times are assumed to be exponential random variables, but this is certainly a simplification. As an example, the stepping process of kinesin is a well-studied though still not completely understood phenomenon. Much work has focused on assessing the dependence of the mean rate of translocation on both the load and the local concentration of adenosine triphosphate (ATP) [43], [37]. Implicit in this analysis is the assumption of exponential wait times with state dependent rate parameters. However, when fitting to data and matching

dispersion information the authors in [12] found it necessary to generalize the wait time distribution. This was followed by more detailed models for which it was shown that load carrying could in fact regularize the stepping times of kinesin motors [36], [8]. Generalizing wait times would significantly affect our results. Since the particle position process is no longer Markov, we no longer have the direct connection to the previous results stated in sections 3.1–3.3, nor can we use the stationary distribution employed in section 4.1 and used for addressing the biological questions in sections 4.2–4.4. In light of the known need for generalized wait times in the stepping process, it seems likely that detailed observation of the rebinding process will call for new mathematical models as well. Recall that when a vesicle unbinds from a microtubule it is unclear whether it typically rebinds to the same microtubule or if it explores the region significantly via diffusion before finding a different microtubule to bind to. In the latter case, a more appropriate model for rebinding time would be to solve some kind of first passage time problem and use that distribution for the rebinding wait.

An important aspect of the biology of axonal transport is not included in the model presented here, namely, the local deposition and eventual degradation of transported materials. For example, sodium channels and sodium pumps are synthesized at the soma, transported down the axon, and deposited in the axonal membrane either uniformly as in an unmyelinated axon or at the nodes of Ranvier in a myelinated axon. Channels and pumps are proteins with half-lives on the order of days to weeks. The present model can be extended to include a deposition compartment at each location, and, clearly, the processes of deposition and subsequent degradation will cause the mean number of particles both on- and off-transport to be monotone decreasing as one moves down the axon. How inhomogeneous this makes the axon will depend on the details of deposition and degradation rates. Our preliminary calculations indicate that long axons, such as the meter-long axons in the human sciatic nerve, would be quite inhomogeneous. This is an important biological issue because it is controversial whether the machinery for protein synthesis (i.e., ribosomes) exist in axons [41]. We have also not included retrograde transport or the fact that some axons may have location-dependent unbinding rates [9]. All of these issues will be the subject of future work.

**Acknowledgments.** The authors are grateful to Professors H. Frederik Nijhout and Vann Bennett of Duke University and Professor Anthony Brown of The Ohio State University for helpful discussions.

#### REFERENCES

- [1] B. ALBERTS, A. JOHNSON, J. LEWIS, M. RAFF, K. ROBERTS, AND P. WALTER, *The Molecular Biology of the Cell*, Garland Science, Taylor & Francis Group, New York, 2008.
- [2] R. D. ALLEN, D. G. WEISS, J. H. HAYDEN, D. T. BROWN, H. FLJIWAKI, AND M. SIMPSON, *Gliding movement and bidirectional transport along single native microtubules from squid axoplasm: Evidence for an active role for microtubules in cytoplasmic transport*, *J. Cell Biol.*, 100 (1985), pp. 1736–1752.
- [3] J. J. BLUM AND M. C. REED, *A model for fast axonal transport*, *Cell Motility*, 5 (1985), pp. 507–527.
- [4] P. C. BRESSLOFF, *Stochastic model of protein receptor trafficking prior to synaptogenesis*, *Phys. Rev. E*, 74 (2006), 031910.
- [5] E. A. BROOKS, *Probabilistic methods for a linear reaction-hyperbolic system with constant coefficients*, *Ann. Appl. Probab.*, 9 (1999), pp. 719–731.
- [6] A. BROWN, *Slow axonal transport: Stop and go traffic in the axon*, *Nat. Rev. Cell Mol. Biol.*, 1 (2000), pp. 153–156.
- [7] N. J. CARTER AND R. A. CROSS, *Mechanics of the kinesin step*, *Nature*, 435 (2005), pp. 308–312.



- [8] R. L. DEVILLE AND E. VANDEN-EIJNDEN, *Regular gaits and optimal velocities for motor proteins*, Biophys. J., 95 (2008), pp. 2681–2691.
- [9] R. DIXIT, J. L. ROSS, Y. E. GOLDMAN, AND E. L. F. HOLZBAUR, *Differential regulation of dynein and kinesin motor proteins by tau*, Science, 319 (2008), pp. 1086–1089.
- [10] S. ETHIER AND T. G. KURTZ, *Markov Processes: Characterization and Convergence*, John Wiley & Sons, New York, 1986.
- [11] J. T. FINER, R. M. SIMMONS, AND J. A. SPUDICH, *Single myosin molecule mechanics: Pico-Newton forces and nanometre steps*, Nature, 368 (1994), pp. 113–119.
- [12] M. E. FISHER AND KOLOMEISKY, *Simple mechanochemistry describes the dynamics of kinesin molecules*, Proc. Natl. Acad. Sci. USA, 98 (2001), pp. 7748–7753.
- [13] A. FRIEDMAN AND G. CRACIUN, *A model of intracellular transport of particles in an axon*, J. Math. Biol., 51 (2005), pp. 217–246.
- [14] A. FRIEDMAN AND G. CRACIUN, *Approximate traveling waves in linear reaction-hyperbolic equations*, SIAM J. Math. Anal., 38 (2006), pp. 741–758.
- [15] A. FRIEDMAN AND B. HU, *Uniform convergence for approximate traveling waves in reaction-hyperbolic systems*, Indiana Univ. Math. J., 56 (2007), pp. 2133–2158.
- [16] A. FRIEDMAN AND B. HU, *Uniform convergence for approximate traveling waves in reaction-diffusion-hyperbolic systems*, Arch. Ration. Mech. Anal., 186 (2007), pp. 251–274.
- [17] A. GENNERICH, A. P. CARTER, S. L. RECK-PETERSON, AND R. D. VALE, *Force-induced bidirectional stepping of cytoplasmic dynein*, Cell, 131 (2007), pp. 952–965.
- [18] R. J. GRIEGO AND R. HERSH, *Random evolutions, Markov chains, and systems of partial differential equations*, Proc. Natl. Acad. Sci. USA, 62 (1969), pp. 305–308.
- [19] G. W. GROSS AND D. G. WEISS, *Theoretical considerations on rapid transport in low viscosity axonal regions*, in Axoplasmic Transport, D. G. Weiss, ed., Springer-Verlag, Berlin, 1982, pp. 330–341.
- [20] H. JASTROW, Electron Microscopic Atlas of Cells, Tissues and Organs on the Internet, Schwann cells section, <http://www.uni-mainz.de/FB/Medizin/Anatomie/workshop/EM/EMSchwannE.html>.
- [21] O. KALLENBERG, *Foundations of Modern Probability*, 2nd ed., Springer-Verlag, New York, 2002.
- [22] R. Z. KHAS'MINSKII, *On stochastic processes defined by differential equations with a small parameter*, Theory Probab. Appl., 11 (1966), pp. 211–228.
- [23] R. Z. KHAS'MINSKII, *A limit theorem for the solutions of differential equations with random right-hand sides*, Theory Probab. Appl., 11 (1966), pp. 390–406.
- [24] F. R. KELLY, *Reversibility and Stochastic Networks*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Chichester, UK, 1979.
- [25] T. G. KURTZ, *Averaging for martingale problems and stochastic approximation*, in Applied Stochastic Analysis, Lecture Notes in Control and Inform. Sci. 177, Springer, Berlin, 1992, pp. 186–209.
- [26] R. J. LASEK AND S. T. BRADY, *The structural hypothesis of axonal transport: Two classes of moving elements*, in Axoplasmic Transport, G. Weiss, ed., Springer-Verlag, Berlin, 1982, pp. 397–405.
- [27] G. LAWLER, *Introduction to Stochastic Processes*, Chapman and Hall, New York, 1995.
- [28] R. MILLER AND R. J. LASEK, *Crossbridges mediate anterograde and retrograde vesicle transport along microtubules in squid axoplasm*, J. Cell Biol., 101 (1985), pp. 2181–2193.
- [29] D. MORAN AND J. C. ROWLEY, III, Visual Histology.com, Chapter 8, Nerves, [http://www.visualhistology.com/products/atlas/VHA\\_Chpt8\\_Nerves.html](http://www.visualhistology.com/products/atlas/VHA_Chpt8_Nerves.html).
- [30] J. M. NEWBY AND P. C. BRESSLOFF, *Directed intermittent search for hidden targets*, New J. Phys., 11 (2009), 023033.
- [31] J. M. NEWBY AND P. C. BRESSLOFF, *Quasi-steady state reduction of molecular motor-based models of directed intermittent search*, Bull. Math. Biol., 72 (2010), pp. 1840–1866.
- [32] S. OCHS, *Rate of fast axoplasmic transport in mammalian nerve fibers*, J. Physiol., 227 (1972), pp. 627–645.
- [33] M. C. REED AND J. J. BLUM, *Theoretical analysis of radioactivity profiles during fast axonal transport: Effects of deposition and turnover*, Cell Motility and the Cytoskeleton, 6 (1986), pp. 620–627.
- [34] M. C. REED AND J. J. BLUM, *Mathematical questions in axonal transport*, in Cell Biology: 1992 Symposium on Some Mathematical Questions in Biology, Lectures in Mathematics in the Life Sciences 24, AMS, Providence, 1994, p. 1.
- [35] M. C. REED, S. VENAKIDES, AND J. J. BLUM, *Approximate traveling waves in linear reaction-hyperbolic equations*, SIAM J. Appl. Math., 50 (1990), pp. 167–180.

- [36] M. J. SCHILSTRA AND S. R. MARTIN, *An elastically tethered viscous load imposes a regular gait on the motion of myosin-V. Simulation of the effect of transient force relaxation on a stochastic process*, J. R. Soc. Interface, 3 (2006), pp. 153–165.
- [37] M. J. SCHNITZER, K. VISSCHER, AND S. M. BLOCK, *Force production by single kinesin motors*, Nat. Cell Biol., 2 (2000), pp. 718–722.
- [38] G. H. STEWART, B. HORWITZ, AND G. W. GROSS, *A chromatographic model of axoplasmic transport*, in Axoplasmic Transport, G. Weiss, ed., Springer-Verlag, Berlin, 1982, pp. 414–422.
- [39] K. SVOBODA, C. SCHMIDT, AND B. SCHNAPP, *Direct observation of kinesin stepping by optical trapping interferometry*, Nature, 365 (1985), pp. 721–727.
- [40] T. TAKENAKA AND H. GOTOH, *Simulation of axoplasmic transport*, J. Theoret. Biol., 107 (1984), pp. 579–601.
- [41] J. L. TWISS AND M. FAIZILBER, *Ribosomes in axons—scrounging from the neighbors*, Trends in Cell Biology, 19 (2009), pp. 236–243.
- [42] R. D. VALE, T. S. REESE, AND M. P. SHEETZ, *Identification of a novel force-generating protein, kinesin, involved in microtubule-based motility*, Cell, 42 (1985), pp. 39–50.
- [43] K. VISSCHER, M. J. SCHNITZER, AND S. M. BLOCK, *Single kinesin molecules studied with a molecular force clamp*, Nature, 400 (1999), pp. 184–189.
- [44] M. A. WELTE AND S. P. GROSS, *Molecular motors: A traffic cop within?*, HFSP J., 2 (2008), pp. 178–182.
- [45] W. D. WILLIS AND R. G. GROSSMAN, *Medical Neurobiology: Neuroanatomical and Neurophysiological Principles Basic to Clinical Neuroscience*, Mosby, St. Louis, MO, 1981.