

A Stochastic Compartmental Model for Fast Axonal Transport

Scott A. McKinley* Lea Popovic† Michael C. Reed‡

October 29, 2018

Abstract

In this paper we develop a probabilistic micro-scale model and use it to study macro-scale properties of axonal transport, the processes by which materials are 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 queueing theory, we can predict macroscopic behavior 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 inhomogeneous does deposition and turnover make the axon?

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

*Department of Mathematics, Duke University, Box 90320, Durham, NC 27701 (mckinley@math.duke.edu)

†Department of Mathematics and Statistics, Concordia University, Montreal QC H3G 1M8

‡Department of Mathematics, Duke University, Box 90320, Durham, NC 27701

[20][24], while slow transport goes at approximately 1 millimeter/day, the rate of axon growth and regeneration [5][20]. The biology and principles of slow transport are not yet clear [5], but the basic mechanisms of fast axonal transport were discovered in the 1980s [2][3][22][34]. 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]. 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 [6][10][16][31]. 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 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 some earlier 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][25][26]. These models did not have traveling wave solutions, but the data certainly looked like approximate traveling waves. In [27] it was shown by a perturbation theory argument that, in the asymptotic limit where the unbinding and binding rates k_1 and k_2 get large, the solution approaches a slowly spreading traveling wave or a normal pulse. Recently, in a series of papers, Friedman and co-workers have introduced new PDE models and proved these results rigorously[12][13][14][15].

Probabilistic models for axonal transport were introduced and used for simulations already in the 1980s [30][32]. However, rigorous work began with Lawler [21] in 1995 and was continued by Brooks [4] 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.

1.1 Summary of Results

In this paper we revise the existing probabilistic models in order to observe randomness in the system as a whole rather than exclusively from the particle's point of view. Our goal is to exploit experimental observations at multiple scales in order to make predictions about the perceived homogeneity of material along the length of the axon. To this end, we propose in Section 2 a continuous-time Markov chain queueing model for the axonal transport system and estimate the order of magnitude of the various parameters. The precise values of the parameters will vary for different molecular motors, for various types of cargo and for various animal species.

In Section 3 we take the individual vesicle point-of-view to recover previous results [4] [26] and show that this model does unify and extend the existing probabilistic and PDE models. Using standard results from renewal theory we calculate the mean velocity and the near-Gaussian wave-front spreading of the law of the vesicle's location. Since we assume independence of the particle interactions, the law of an individual is equivalent to the distribution of an ensemble of particles released at the same time. Therefore the PDE governing a spatial-continuum limit of the law of a single particle is the same as the PDE studied in [26] (see Section 3.3).

Subsequently 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. Via the coefficient of variation we give a precise characterization in Section 4.2 of the experimenter's qualitative perception of homogeneity at the millimeter scale.

The final two sections deal with model predictions. In Section 5 we study the non-equilibrium dynamics of recovery. Due to the product structure of the law of the transient dynamics, 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 lengthscale of interest. In Section 6 we consider the effect of deposition of material in the cell membrane. We find that there should be an exponential loss of material along the length of the cell, and conclude that a new biological mechanism must be discovered to account for the homogeneity that is a hallmark of axonal transport experimental observations.

2 The model and its parameters

Let L be the length of the axon, divided evenly into N lateral sections each of length δ , 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 with mean lateral velocity v , or
- an off-transport state with lateral velocity 0.

We use a $2N$ -dimensional continuous time Markov chain to model the particle dynamics:

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

Each Q_i and P_i has the natural numbers $\mathbb{N} = \{0, 1, 2, \dots\}$ as its state space. In the simplest model, we consider the following transitions and rates,

- *Lateral transport:*
 $(Q_i, Q_{i+1}) \rightarrow (Q_i - 1, Q_{i+1} + 1)$ at rate $rQ_i(t)$, where $r = v/\delta$;
- *Switch from on-transport to off-transport:*
 $(Q_i, P_i) \rightarrow (Q_i - 1, P_i + 1)$ at rate $k_1Q_i(t)$;
- *Switch from off-transport to on-transport:*
 $(P_i, Q_i) \rightarrow (P_i - 1, Q_i + 1)$ at rate $k_2P_i(t)$;
- *Production of new particles:*
 $(Q_1) \rightarrow (Q_1 + 1)$ at rate q_0r ;
- *Removal of particles at distal end:*
 $(Q_N) \rightarrow (Q_N - 1)$ at rate $rQ_N(t)$.

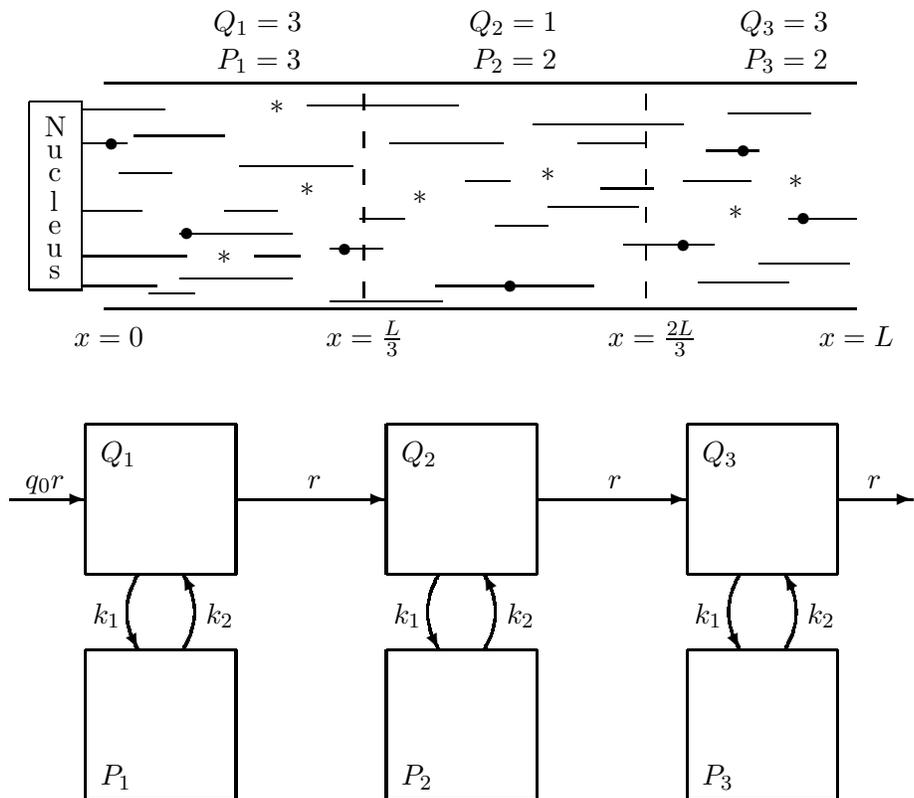


Figure 1: Double Chain

The lateral transport rate, r , is inversely proportional to the length scale so that the mean number of particles per unit length is invariant with respect to rescaling δ . The rate of production, $q_0 r$, ensures that this mean number per unit scales with q_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.

This is an exactly solvable model. 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 10^{-6} m/s . We have already stated the assumption that the length scale of the individual steps satisfies $\delta \sim 10^{-8}m$. This implies that the rate parameter should be $r = v/\delta = 10^2 s^{-1}$.

We now turn our attention to the on-off rates k_1 and k_2 . 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_1 . Within the model, at each step on the transport mechanism the particle has a binary decision to jump laterally along the transport with probability $r/(k_1 + r)$, or to jump off with probability $k_1/(r + k_1)$. The number of jumps along the transport system before jumping off is therefore geometrically distributed on the set $\{0, 1, \dots\}$ with success probability $\frac{r}{r+k_1}$. It follows that average number of steps in the run is $\frac{r}{k_1}$ and therefore the average run length is $\frac{r}{k_1} \times 10^{-8} m$. Setting this equal to the average experimental run length of 10^{-6} from [9], we see that $\frac{r}{k_1} \sim 10^2$, implying that $k_1 \sim 1s^{-1}$. 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 $\frac{k_1}{k_2}$. Dixit [9] found that approximately 75% of the particles were motile so this ratio is approximately equal to 3. Since $k_1 \sim 1s^{-1}$ we see that $k_2 \sim \frac{1}{3}$.

It remains to estimate q_0 . We will see in Proposition 4.1 that the mean number of particles per unit length is $(1 + \frac{k_1}{k_2})q_0 = 4q_0$. 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 [17], Fig. 3; [18]; [23]), 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 “box” 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 standard theorems from queuing theory. In the $\delta \rightarrow 0$ limit, the law of the location of a single particle converges to the Green’s function of a linear partial differential equation. This enables us to obtain, as a special case, the asymptotic behavior of the PDE models for axonal transport in various asymptotic limits.

3.1 The active transport mode

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

Proposition 3.1. *Let $k_1 = k_2 = 0$, and $r = v/\delta > 0$, then $X_t \sim \text{Pois}(vt)$. In particular, the mean velocity of the particle is given by $\frac{1}{t}\mathbb{E}[X_t] = v$. For any given $t \geq 0$, in the limit as $\delta \rightarrow 0$ the position of the particle satisfies*

$$\frac{X_t}{t} \xrightarrow{\text{a.s.}} v \quad \text{and} \quad \frac{1}{\sqrt{\delta}}(X_t - vt)_{t \geq 0} \xrightarrow{d} \sqrt{v}(B_t)_{t \geq 0}$$

where B is a standard Brownian motion.

Proof. Since $X_t = \delta n_t$ where $(n_t)_{t \geq 0}$ is a Poisson process with rate $r = v/\delta$ the set out results follow from the law of large numbers (LLN) and the central limit theorem (CLT) for a Poisson process. \square

3.2 The on/off dynamics

We now consider a particle which undergoes transitions from on-transport to off-transport state and back. Denote again by $X_t \in \{0, \delta, 2\delta, \dots, L\}$ the lateral position of a particle at time t and let n_t 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.

Proposition 3.2. *Let $k_1, k_2 > 0$, and let $r = v/\delta > 0$, then the mean velocity of the particle satisfies $\frac{1}{t}\mathbb{E}[X_t] \xrightarrow[t \rightarrow \infty]{} \frac{k_2}{k_1+k_2}v$.*

Proof. Consider the time τ for a particle to make one step of lateral transport. Before doing so a particle performs m switches from on-transport to off-transport and back, where m is distributed as a Geometric($\frac{r}{k_1+r}$) variable. The amounts of time a particle takes for each switch $\tau_i^{Q \rightarrow P}$ from on-transport to off-transport are iid variables with Exponential(k_1+r) distribution. Likewise, the amounts of time a particle takes for each switch $\tau_i^{P \rightarrow Q}$ from off-transport back to on-transport are iid variables with Exponential(k_2) distribution, and are independent of the other switching times. Then

$$\tau = \sum_{i=1}^m (\tau_i^{Q \rightarrow P} + \tau_i^{P \rightarrow Q}) + \tau_{m+1}^{Q \rightarrow Q}$$

with $\mathbb{E}[\tau] = (k_1 + k_2)/(k_2 r)$. Let n_t be the number of times a particle makes a step of lateral transport until time t . By the Renewal Theorem $\frac{1}{t}\mathbb{E}[n_t] \xrightarrow[t \rightarrow \infty]{} \frac{1}{\mathbb{E}[\tau]}$, and the result follows from $X_t = \delta n_t$ and $\delta \frac{k_2 r}{k_1+k_2} = \frac{k_2 v}{k_1+k_2}$. \square

Proposition 3.3. *Let $k_1, k_2 > 0$, and $r = v/\delta > 0$. In the limit as $\delta \rightarrow 0$ the position of a particle satisfies*

$$\frac{X_t}{t} \xrightarrow[t \rightarrow \infty]{a.s.} \frac{k_2}{k_1+k_2}v, \quad \text{and} \quad \sqrt{t} \left(\frac{X_t}{t} - \frac{k_2}{k_1+k_2}v \right) \xrightarrow[t \rightarrow \infty]{d} \sqrt{\frac{2k_1 k_2}{(k_1+k_2)^3}} v B_1$$

where $B_1 \sim \text{Normal}(0, 1)$.

Proof. The number of lateral transport steps $(n_t)_{t \geq 0}$ is a renewal chain with $\mathbb{E}[\tau] = (k_1 + k_2)/k_2 r$ and $\text{Var}[\tau] = ((k_1 + k_2)^2 + 2k_1 r)/r^2 k_2^2$. By the LLN and CLT for Renewal chains

$$\frac{n_t}{t} \xrightarrow[t \rightarrow \infty]{a.s.} \frac{1}{\mathbb{E}[\tau]}, \quad \text{and} \quad \sqrt{t} \left(\frac{n_t}{t} - \frac{1}{\mathbb{E}[\tau]} \right) \xrightarrow[t \rightarrow \infty]{d} \sqrt{\frac{\text{Var}[\tau]}{(\mathbb{E}[\tau])^3}} B_1$$

and now the result follows from $X_t = \delta n_t$, $\delta \frac{1}{\mathbb{E}[\tau]} = \delta \frac{k_2 r}{k_1 + k_2} = \frac{k_2}{k_1 + k_2} v$, and

$$\delta^2 \frac{\text{Var}[\tau]}{(\mathbb{E}[\tau])^3} = \frac{\delta k_2 v}{k_1 + k_2} + \frac{2k_1 k_2 v^2}{(k_1 + k_2)^3} \approx \frac{2k_1 k_2}{(k_1 + k_2)^3} v^2 \text{ for } \delta \approx 0.$$

□

The fact that an individual particle starting at $x = 0$ will have the distribution given by Proposition 3.3 at long times was also obtained by Brooks [4] who proved tail estimates on the CLT and found that the remainder is $O(\frac{1}{\sqrt{t}})$. In our approach, one can obtain the same error estimate using the asymptotic analysis of renewal chains, see, for example, [7].

3.3 Connection to Partial Differential Equations Models.

In order to demonstrate the connection between our model and the PDEs seen in [27][26][14], we make a formal calculation regarding convergence of the law of the location of a single particle in the full system.

Let $\delta_N = \frac{1}{N}$ and let $X_N(t)$ denote the lateral position of a particle in a system with N boxes. Define

$$\begin{aligned} q_N(x, t) &:= \mathbb{P}\{X_N(t) = x, \text{ and the particle is } \textit{on} \text{ track}\} \\ p_N(x, t) &:= \mathbb{P}\{X_N(t) = x, \text{ and the particle is } \textit{off} \text{ track.}\} \end{aligned}$$

We suppress N in the notation. By definition of the Poisson process, for small values of h we have

$$\begin{aligned} q(x, t + h) &= q(x, t)(1 - (k_1 + r)h) + q(x - \delta, t)rh + p(x, t)k_2h + o(h) \\ p(x, t + h) &= p(x, t)(1 - k_2h) + q(x, t)k_1h + o(h). \end{aligned}$$

which we may reorganize as

$$\begin{aligned} \frac{1}{h}(q(x, t + h) - q(x, t)) &= r[q(x - \delta, t) - q(x, t)] - k_1q(x, t) + p(x, t)k_2 + \frac{o(h)}{h} \\ \frac{1}{h}(p(x, t + h) - p(x, t)) &= -k_2p(x, t) + k_1q(x, t) + \frac{o(h)}{h}. \end{aligned}$$

Formally, we take the Taylor expansion of q in space: $q(x - \delta, t) = q(x, t) - \delta \partial_x q(x, t) + o(\delta^2)$. Taking limits as $h \rightarrow 0$ and $\delta \rightarrow 0$ yields the system of PDEs

$$\partial_t q(x, t) + v \partial_x q(x, t) = -k_1 q(x, t) + k_2 p(x, t) \quad (1)$$

$$\partial_t p(x, t) = k_1 q(x, t) - k_2 p(x, t). \quad (2)$$

When $k_1 = 0 = k_2$, 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 translation of the delta function, while the time evolution via the equations (1) and (2) will have a spreading profile. This is seen in Propositions 3.1 and 3.3 since the variance of the fluctuations goes to 0 in Proposition 3.1 but not in Proposition 3.3 as $\delta \rightarrow 0$.

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 wave front moving at constant velocity away from the cell body. The equations (3) and (4) are linear and do not have solutions that are bounded traveling waves. However, it was shown by a perturbation theory argument in [27][26] that as $\varepsilon \rightarrow 0$ the solutions of

$$\varepsilon(\partial_t + v\partial_x)q_\varepsilon(x, t) = -k_1q_\varepsilon(x, t) + k_2p_\varepsilon(x, t), \quad q_\varepsilon(0, t) = q_0 \quad (3)$$

$$\varepsilon\partial_t p_\varepsilon(x, t) = k_1q_\varepsilon(x, t) - k_2p_\varepsilon(x, t). \quad (4)$$

are to leading order

$$q_\varepsilon(x, t) = H\left(\frac{x - at}{\varepsilon^{1/2}}, t\right), \quad p_\varepsilon(x, t) = cH\left(\frac{x - at}{\varepsilon^{1/2}}, t\right),$$

where H satisfies the heat equation

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

and

$$a = \frac{k_1v}{k_1 + k_2}, \quad \kappa^2 = \frac{2k_1k_2v^2}{(k_1 + k_2)^3}. \quad (5)$$

This asymptotic form is valid for small ε , that is for large k_1 and k_2 . 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) and (4), so the solutions of (3) and (4) behave like approximate traveling waves for large t and large x whether or not k_1 and k_2 are large. The results suggested by these perturbation theory arguments have been proven rigorously by Friedman and coworkers [12][13][14][15].

In the case that radiolabeled particles enter the axon only for a short time, $0 \leq t \leq T$, the same perturbation analysis shows that for large times the solution of (3) and (4), where $q(0, t) = 0$ for $t > T$ is to leading order a spreading Gaussian with mean and variance as given in Proposition 3.3. This

corresponds to what is seen experimentally and arises asymptotically in the solutions of (3) and (4) by convolving the asymptotic solution corresponding to the initial condition $q(x, 0) = \delta_0(x)$ with the indicator function $\chi_{[0, T]}(x)$.

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. *If the incoming rate of particles at Q_0 is $q_0 r$, then the system has the following stationary distribution*

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

where the $\{Q_i\}$ and $\{P_i\}$ are mutually independent.

Proof. Given the value of the input rate $Q_0 = q_*$ the generator of the process (Q_1, P_1) in the first section is

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

If we first use $f(q, p) = Q_1(t)$ then use $f(q, p) = P_1(t)$ and take expectations we get a system of ODE's governing the change in $\mathbb{E}[Q_1], \mathbb{E}[P_1]$ over time

$$\begin{aligned} \frac{d\mathbb{E}[Q_1](t)}{dt} &= r\mathbb{E}[Q_0] - r\mathbb{E}[Q_1(t)] - k_1\mathbb{E}[Q_1(t)] + k_2\mathbb{E}[P_1(t)] \\ \frac{d\mathbb{E}[P_1](t)}{dt} &= k_1\mathbb{E}[Q_1(t)] - k_2\mathbb{E}[P_1(t)] \end{aligned}$$

indicating that in equilibrium $\mathbb{E}[Q_1] = \mathbb{E}[Q_0]$, $\mathbb{E}[P_1] = \mathbb{E}[Q_1] \frac{k_1}{k_2}$. For purposes that will soon be clear we first assume that Q_0 has a constant $\text{Pois}(\lambda_*)$ distribution over time rather than just being equal to λ_* . Let $\pi(q_*, q, p) = \pi_{\lambda_*}(q_*) \otimes \pi_{\lambda_Q}(q) \otimes \pi_{\lambda_P}(p)$ where π_λ are $\text{Pois}(\lambda)$ distributions with rates $\lambda_*, \lambda_Q = \lambda_*$, and $\lambda_P = \lambda_* \frac{k_1}{k_2}$ respectively. To show that $\pi(q_*, q, p)$ is a stationary distribution for the process (Q_1, P_1) we need to check that

$$\sum_{q_*=0}^{\infty} \sum_{q=0}^{\infty} \sum_{p=0}^{\infty} \mathcal{A}_{q_*} f(q, p) \pi(q_*, q, p) = 0$$

for any choice of function $f \in \mathcal{D}(\mathcal{A}_{q_*})$.

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

In the inner two sums, for a fixed value of q_* , the factor multiplying $f(q, p)$ for any $q, p \in \mathbb{N}^2$ comes only from terms involving $\{q-1, q, q+1\}$ and $\{p-1, p, p+1\}$ and equals $e^{-(\lambda_* + \lambda_Q + \lambda_P)}$ times

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

since $\lambda_Q = \lambda_*$ and $\frac{\lambda_P}{\lambda_Q} = \frac{k_1}{k_2}$.

Summing over q_* the factor multiplying $f(q, p)$ becomes

$$\begin{aligned}
& e^{-(\lambda_Q + \lambda_P)} \frac{\lambda_Q^q}{q!} \frac{\lambda_P^p}{p!} \sum_{q_*=0}^{\infty} e^{-\lambda_*} \frac{\lambda_*^{q_*}}{q_*!} \left(\frac{q}{\lambda_*} r q_* - r q_* + \lambda_* r - r q \right) \\
&= e^{-(\lambda_Q + \lambda_P)} \frac{\lambda_Q^q}{q!} \frac{\lambda_P^p}{p!} \left(q r - r \lambda_* + \lambda_* r - r q \right) = 0
\end{aligned}$$

We've just shown that if the distribution of input particles Q_0 in equilibrium is $\text{Pois}(\lambda_*)$ then the stationary distribution of Q_1, P_1 is independent of the distribution of Q_0 and is $\pi_{\lambda_Q} \otimes \pi_{\lambda_P}$ where $\lambda_Q = \mathbb{E}[Q_0]$, and $\lambda_P = \mathbb{E}[Q_0]k_1/k_2$. It is easy to see in the above calculation that if the input Q_0 was indeed constant q_0 then the stationary distribution of (P_1, Q_1) would again be $\pi_{\lambda_Q} \otimes \pi_{\lambda_P}$ with rates $\lambda_Q = q_0$ and $\lambda_P = q_0k_1/k_2$ respectively. Now, since the stationary distribution of Q_1 is the input into (Q_2, P_2) process, the more general calculation shows that in equilibrium the stationary distribution of (Q_2, P_2) is independent of the stationary distribution for Q_1 and is $\pi_{\lambda_Q} \otimes \pi_{\lambda_P}$ with rates $\lambda_Q = \mathbb{E}[Q_1] = q_0$ and $\lambda_P = \mathbb{E}[Q_1]k_1/k_2 = q_0k_1/k_2$ again. It follows by induction that the stationary distributions for $\{(Q_i, P_i)\}$ are independent and identically distributed as $\pi_{\lambda_Q} \otimes \pi_{\lambda_P}$, $\lambda_Q = q_0$, $\lambda_P = q_0k_1/k_2$. This is also an example of a clustering process satisfying the detailed balance conditions with linear rates discussed in Sec. 8.2 of [19].

□

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 $\nu = 1, 2, \dots 8$. Let $\Delta = 10^\nu \delta$; we want to determine how similar different segments of the axon of size Δ are. Let Q_Δ and P_Δ denote the numbers of on-track and off-track particles in a segment of length Δ .

In equilibrium, Q_Δ and P_Δ are both sums of 10^ν independent Poisson random variables with parameters $\lambda_Q = q_0$ and $\lambda_P = \frac{k_1}{k_2}q_0$, respectively. Therefore the distributions of Q_Δ and P_Δ 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 Δ is $10^\nu(\lambda_Q + \lambda_P)$. To see how homogeneous different slices of length Δ are, we consider the coefficient of variation, c_Δ , which is the standard deviation divided by the mean.

$$c_\Delta = \frac{1}{\sqrt{(\lambda_Q + \lambda_P)10^\nu}} = \frac{1}{\sqrt{(1 + k_1/k_2)q_0\Delta 10^8}}.$$

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 $k_1/k_2 = 3$, 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.

5 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 the system of transport with reversible binding is that if it is locally out of equilibrium, the dynamics will automatically take it back to equilibrium. 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? We study this question first for a single location and then use the estimates derived to scale the results to segments of any length.

Proposition 5.1. *Let $(Q, P)(0) = (0, 0)$ and let the constants $a > 0$ and $p \in (0, 1)$ satisfy the relationship $a^2 p |\lambda_\infty| > 1$ where λ_∞ is the equilibrium vector of $(Q, P)(t)$. Then there exists a t_* such that $t \geq t_*$,*

$$\mathbb{P}\{|(Q, P)(t) - \lambda_\infty| \geq a|\lambda_\infty|\} \leq p.$$

In fact, the choice

$$t_* = \alpha^{-1} \ln \left(\frac{\sqrt{p|\lambda_\infty|}}{a\sqrt{p|\lambda_\infty|} - 1} \right)$$

is sufficient, where

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

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

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

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

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

Since the initial condition for both P and Q are less than their respective equilibrium values, each are monotonically increasing functions and the above reduces to

$$\mathbb{P}\{|(Q, P)(t) - \boldsymbol{\lambda}_\infty| \geq a|\boldsymbol{\lambda}_\infty|\} \leq \frac{1}{a^2(1-\beta)^2|\boldsymbol{\lambda}_\infty|}$$

for all $t > t_*$. In order to satisfy the requirement that the right hand side must be less than p , we solve for β and find

$$\beta = 1 - \frac{1}{a\sqrt{p|\boldsymbol{\lambda}_\infty|}}$$

provided that $a\sqrt{p|\boldsymbol{\lambda}_\infty|} > 1$.

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

$$\dot{\boldsymbol{\lambda}}(t) = -A_1\boldsymbol{\lambda}(t) + q_0r\mathbf{e}_1 \quad (6)$$

where \mathbf{e}_1 is the unit vector $(1, 0)$ and

$$A_1 = \begin{pmatrix} k_1 + r & -k_2 \\ -k_1 & k_2 \end{pmatrix}.$$

The solution to (6) is

$$\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}_\infty + e^{-A_1t}(\boldsymbol{\lambda}(0) - \boldsymbol{\lambda}_\infty)$$

where $\boldsymbol{\lambda}_\infty = q_0rA_1^{-1}\mathbf{e}_1 = q_0(1, \frac{rk_1}{k_2})$. This yields the estimate

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

where α is the smaller of the eigenvalues of A_1 , namely

$$\alpha = \frac{1}{2}(k_1 + k_2 + r - \sqrt{(k_1 + k_2 + r)^2 - 4k_2r}).$$

Noting that $\alpha > 0$, t_* may be chosen so that $e^{-\alpha t_*} \leq a\beta$, i.e.

$$t_* = \alpha^{-1} \ln(1/(a\beta))$$

□

We now suppose that the whole axon is in statistical equilibrium except for a segment of length $\delta \times 10^\nu m$ in which we will assume that there are no particles either on or off the track. We want to know how long will it take for this segment to get back close to equilibrium. Of course, Proposition 5.1 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 $\delta \times 10^\nu m$. In this rescaled system, the unbinding and binding rates per particle, k_1 and k_2 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 $\tilde{r} = r \times 10^{-\nu}$.

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

$$\frac{d}{dt} \tilde{\lambda}(t) = -\tilde{A}_1 \tilde{\lambda}(t) + q_0 r \mathbf{e}_1 \quad (7)$$

with

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

We note that the last term in (7) 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,

$$\tilde{\lambda}_\infty = q_0 r \tilde{A}_1^{-1} \mathbf{e}_1 = q_0 \frac{r}{\tilde{r}} \begin{pmatrix} 1 \\ k_1/k_2 \end{pmatrix}.$$

Both components of this vector are of order 10^ν , as expected.

We will use the parameters discussed in Section 2: $k_1 = 1, k_2 = \frac{1}{3}, v = 10^{-6} m/s, r = 10^2 s^{-1}$. This implies $\tilde{r} = 10^{2-\nu} s^{-1}$. As in our analysis in Proposition 5.1, the time to equilibrium, \tilde{t}_* , is proportional to $\tilde{\alpha}^{-1}$ where $\tilde{\alpha}$ satisfies:

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

We are most interested in the cases $\nu \geq 3$ (the segment has length ≥ 10 microns). Then $k_2\tilde{r}$ is small compared to k_1 . Ignoring constants and restricting our attention to the leading order terms gives

$$\tilde{\alpha} \sim \frac{k_2\tilde{r}}{k_1 + k_2 + \tilde{r}} \sim 10^{2-\nu} s^{-1}$$

Thus,

$$\tilde{t}_* \sim 10^{\nu-2} s.$$

Thus, for a 10 micron segment ($\nu = 3$) the time to equilibrium is about 10 seconds and for a 1 millimeter segment ($\nu = 5$) the time to equilibrium is 1000 seconds or 15 minutes. Of course, the time to get close to equilibrium depends on the parameter p that represents what we mean by “close.”

6 Transport with Deposition

One major feature of the biology we have ignored is that some particles in the off-transport state are actually deposited permanently in the membrane or are used for some other purpose. Such a phenomenon is easy to add to the model proposed above, but we must address a new qualitative feature of the results. Namely there is an exponential loss of material as we move toward the distal end.

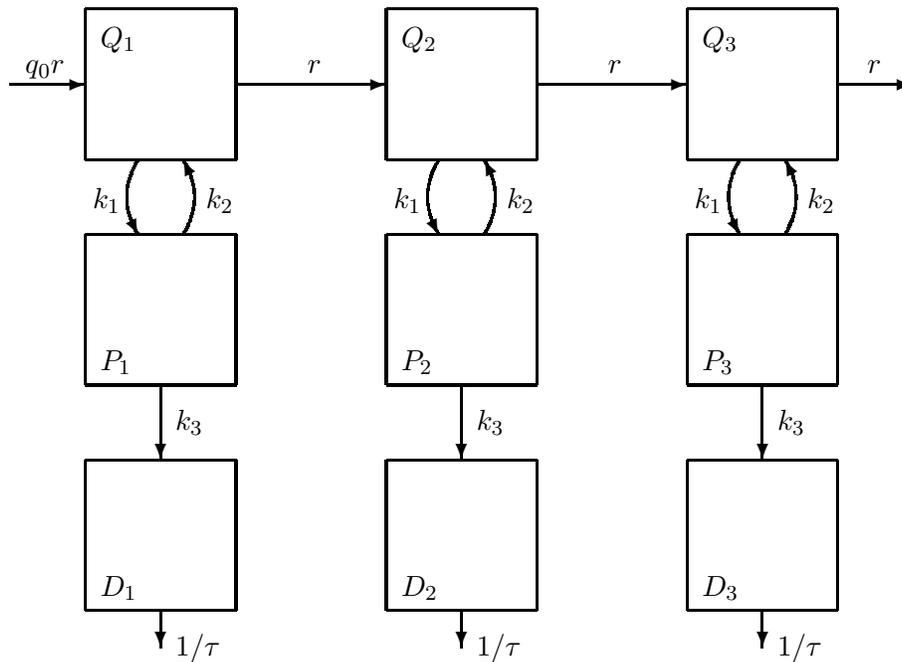


Figure 2: Fast Axonal Transport with Deposition

In addition to deposition, we recognize that particles are only useful in the membrane for a random amount of time with a given half-life, after which they decompose and get transported back to the cell body. Neglecting this retrograde transport, we augment the previous model as follows: at each δ -length lateral section let $D_i(t)$ be the number of particles deposited in section i , with $\mathbb{N} = \{0, 1, 2, \dots\}$ as the state space for D_i . In addition to the previous dynamics we also define the following transition rates

- $(P_i, D_i) \rightarrow (P_i - 1, D_i + 1)$ (deposition) at rate $k_3 P_i(t)$
- $D_i \rightarrow D_i - 1$ (decay) at rate $1/\tau$ where $\tau > 0$

The constant τ in the decay rate is the average lifetime of a deposited particle.

6.1 Dynamics from the particle perspective

As before, we can use the renewal theorem to determine the average speed of particles that are *not* deposited, as well as the spreading of the wavefront.

Proposition 6.1. *Let $k_1, k_2, k_3 > 0$, $\tau < \infty$ and let $r = v/\delta > 0$, then the mean velocity of the particle that is not deposited into the membrane by time t satisfies $\frac{1}{t}\mathbb{E}[X_t] \xrightarrow[t \rightarrow \infty]{} \frac{k_2 + k_3}{k_1 + k_2 + k_3}v$.*

Moreover, in the limit as $\delta \rightarrow 0$ the position of a particle satisfies

$$\frac{X_t}{t} \xrightarrow[t \rightarrow \infty]{a.s.} \frac{k_2 + k_3}{k_1 + k_2 + k_3}v, \quad \sqrt{t} \left(\frac{X_t}{t} - \frac{k_2 + k_3}{k_1 + k_2 + k_3}v \right) \xrightarrow[t \rightarrow \infty]{d} \sqrt{\frac{2k_1(k_2 + k_3)}{(k_1 + k_2 + k_3)^3}}v B_1$$

where $B_1 \sim \text{Normal}(0, 1)$.

Remark 1. Notice the mean velocity of particles is higher than when there is no deposition

$$\frac{k_2 + k_3}{k_1 + k_2 + k_3}v \geq \frac{k_2}{k_1 + k_2}v, \text{ with equality iff } k_3 = 0$$

A simple explanation for this is that particles that are still in lateral transport had returned to on-transport back from off-transport before managing to be deposited into the membrane, meaning that their time off-transport was shorter than it would have been if they did not have to beat the exponential clock for deposition.

Proof. When there is deposition then by time t each particle has either already deposited, or is still performing lateral transport. A particle that is

deposited into the membrane has velocity zero from the time of its deposition on. A particle that can still be laterally transported must have avoided deposition until time t .

If we proceed along the lines of the proof without deposition, then the only difference now is that the amount of time a particle takes for each switch from off-transport back to on-transport are i.i.d. exponential variables with parameter k_2+k_3 rather than k_2 . Hence the amount of time τ for a particle to make a step from one section to another has $\mathbb{E}[\tau] = (k_1+k_2+k_3)/((k_2+k_3)r)$, and by the Renewal Theorem the number of steps a particle that can still be laterally transported satisfies $\frac{1}{t}\mathbb{E}[n_t] \xrightarrow{t \rightarrow \infty} \frac{(k_2+k_3)}{(k_1+k_2+k_3)}r$.

Also $\text{Var}[\tau] = ((k_1+k_2+k_3)(k_1+k_2+k_3+2k_1r))/r^2(k_2+k_3)^2$, so the result now follows by the Renewal LLN and CLT for n_t and the fact that

$$\delta \frac{1}{\mathbb{E}[\tau]} = \delta \frac{(k_2+k_3)r}{k_1+k_2+k_3} = \frac{k_2+k_3}{k_1+k_2+k_3}v,$$

and

$$\delta^2 \frac{\text{Var}[\tau]}{(\mathbb{E}[\tau])^3} = \frac{\delta(k_2+k_3)v}{k_1+k_2+k_3} + \frac{2k_1(k_2+k_3)v^2}{(k_1+k_2+k_3)^3} \approx \frac{2k_1(k_2+k_3)}{(k_1+k_2+k_3)^3}v^2 \quad \text{for } \delta \approx 0$$

□

6.2 Dynamics from the spatial system perspective

Once again we compute the stationary distribution of the flow-through system, this time with deposition.

Proposition 6.2. *Suppose the incoming rate of particles at Q_0 is q_0r . Then the stationary distribution of the system is*

$$Q_i \sim \text{Pois}(q_0 \rho^i), \quad P_i \sim \text{Pois}\left(\frac{k_1}{k_2+k_3}q_0 \rho^i\right), \quad D_i \sim \text{Pois}\left(\frac{k_1 k_3 \tau}{k_2+k_3}q_0 \rho^i\right)$$

where $\rho = r(k_2+k_3)/((r+k_1)(k_2+k_3)-k_1k_2)$ and $\{Q_i\}$, $\{P_i\}$, and $\{D_i\}$ are mutually independent.

Proof. Since (Q_i, P_i) do not depend on D_i the proof for the stationary distribution of (Q_i, P_i) follows along the same lines as in the case when there is no deposition. We first check the stationary distribution for the process (Q_1, P_1) in the first section. For given value of $Q_0 = q_*$ the generator of the

process (Q_1, P_1) is given by

$$\begin{aligned} \mathcal{A}_{q_*} f(Q_1, P_1) &= [f(Q_1 + 1, P_1) - f(Q_1, P_1)]r q_* + [f(Q_1 - 1, P_1) - f(Q_1, P_1)]r Q_1 \\ &\quad + [f(Q_1 - 1, P_1 + 1) - f(Q_1, P_1)]k_1 Q_1 + [f(Q_1 + 1, P_1 - 1) - f(Q_1, P_1)]k_2 P_1 \\ &\quad + [f(Q_1, P_1 - 1) - f(Q_1, P_1)]k_3 P_1 \end{aligned}$$

The system of ODE's governing the change in $\mathbb{E}[Q_1], \mathbb{E}[P_1]$ is now

$$\begin{aligned} \frac{\partial_t \mathbb{E}[Q_1]}{\partial t} &= \mathbb{E}[Q_0]r - \mathbb{E}[Q_1]r - \mathbb{E}[Q_1]k_1 + \mathbb{E}[P_1]k_2 \\ \frac{\partial_t \mathbb{E}[P_1]}{\partial t} &= \mathbb{E}[Q_1]k_1 - \mathbb{E}[P_1]k_2 - \mathbb{E}[P_1]k_3 \end{aligned}$$

Hence, in equilibrium

$$\mathbb{E}[Q_1] = \mathbb{E}[Q_0] \frac{k_2 + k_3}{(r + k_1)(k_2 + k_3) - k_1 k_2} = q_0 \rho$$

and

$$\mathbb{E}[P_1] = \mathbb{E}[Q_1] \frac{k_1}{k_2 + k_3} = \frac{k_1 q_0}{k_2 + k_3} \rho.$$

If we let $\lambda_Q = \mathbb{E}[Q_1], \lambda_P = \mathbb{E}[P_1]$, and define the distribution $\pi(Q_0, Q_1, P_1) = \pi_{q_0 \delta}(Q_0) \otimes \pi_{\lambda_Q}(Q_1) \otimes \pi_{\lambda_P}(P_1)$, then $\sum_{q_*, q, p=0}^{\infty} \mathcal{A}_{q_*} f(q, p) \pi(q_*, q, p) = 0$ can be verified exactly as before. Since now the mean of Q_1 differs from the mean of Q_0 by a multiplicative factor $\rho < 1$, and this is the mean of the input into Q_2 , it follows by induction that the sequence of rates for $\{Q_i, P_i\}$ Poisson random variables decays geometrically in i by a factor ρ^i . This immediately implies that in equilibrium the magnitude of the variances also decays geometrically along the line of progression of lateral transport.

To find the stationary distribution for D_i note that if the number of particles at $P_i = p_*$ the generator for D_i is

$$A_{p_*} f(d) = [f(d + 1) - f(d)]k_3 p_* + [f(d - 1) - f(d)](1/\tau)d$$

It is easily checked that if $P_i \sim \pi_{\lambda_{P_i}}$ then $\pi_{\lambda_{D_i}}$ is the stationary distribution for D_i which is independent of the distribution for P_i with $\lambda_{D_i} = k_3 \tau \lambda_{P_i}$. \square

Corollary 6.3. *When in equilibrium the mean and the variance of the number of particles along the axon decay exponentially, with a loss ratio*

$$\frac{\mathbb{E}[Q_{i+1}]}{\mathbb{E}[Q_i]} = \frac{\text{Var}[Q_{i+1}]}{\text{Var}[Q_i]} \approx \exp\left(-\frac{\delta k_1}{k_2 + k_3} \frac{k_3}{v}\right) \quad (8)$$

and the same loss ratio holds for the off-transport and the deposited particles.

Remark 2. Notice that the rate of loss depends on the ratio of reaction rates k_1/k_2 , which we earlier estimated to be approximately 3, as well as the ratio k_3/v , and that if $k_1, k_2 \gg k_3$ then the loss rate is $e^{-\frac{3k_3}{v}}$.

Proof. Recalling $\rho = r(k_2 + k_3)/(r(k_2 + k_3) + k_1k_3)$ and $r = v/\delta$, we rewrite

$$\rho = \left(1 + \delta \frac{k_1k_3}{v(k_2 + k_3)}\right)^{-1}$$

Observing that for any small δ , the approximation $(1 + \delta x)^{-1} = 1 - \delta x + (\delta x)^2 - \dots \approx e^{-\delta x}$, we have Equation (8).

The same ratio holds for the other two states as well

$$\mathbb{E}[P_i]/\mathbb{E}[P_0] = \text{Var}[P_i]/\text{Var}[P_0] = \mathbb{E}[D_i]/\mathbb{E}[D_0] = \text{Var}[D_i]/\text{Var}[D_0] = \rho^i$$

□

6.3 Homogeneity despite deposition

It remains to estimate the rates k_3 and τ from experimental data and then make a formal calculation to address the volume of material loss due to deposition and turnover as a function of length scale. To be precise, suppose we want to know the percentage loss over a length $\Delta = 10^\nu \delta$. Then from the Corollary 6.3 the fraction of material loss is approximately

$$\frac{\mathbb{E}[Q_{10^\nu}]}{\mathbb{E}[Q_1]} \approx \prod_{i=1}^{10^\nu} \exp\left(-\delta \frac{k_1}{k_2 + k_3} \frac{k_3}{v}\right) = \exp\left(-10^{\nu-8} \frac{k_1}{k_2 + k_3} \frac{k_3}{v}\right) \quad (9)$$

So, for example, at the meter scale, $\nu = 8$, and so the loss ratio is $\exp(-k_1k_3/v(k_2 + k_3))$.

We now consider the parameters of a specific example, sodium pumps being deposited in the cell membrane. We assume that the half life of a single pump is approximately a week, which means $\tau = 6 \times 10^5 s$. It remains to estimate k_3 . According to [36], there are roughly 1000 sodium pumps per micron² of membrane surface area. If we take the diameter of the neuron to be 5 microns, then the membrane surface area for one 10 nm slice is about 0.15 μm^2 . So there are roughly 150 sodium pumps per slice.

It is important to note that each vesicle carries roughly 100 sodium pumps and so the stationary distribution for the quantity D_i is on average 1/100th of the number of sodium channels which are deposited into the membrane at the location i . Taking this into account and recalling our earlier

parameter estimations, $k_1 = 1$, $k_2 = 1/3$, $v = 10^{-6}$ and $q_0 = 1$ we estimate the average number of sodium pumps in the membrane of the first 10 *nm* slice to satisfy $150 = 100 q_0 k_1 k_3 \tau / (k_2 + k_3)$ which implies $k_3 = 1 \times 10^{-6}$.

From the above we see that the loss ratio at the meter scale is

$$\exp\left(-\frac{k_1}{k_2 + k_3} \frac{k_3}{v}\right) = e^{-3} \approx 0.04 \quad (10)$$

which is to say that over the length of an entire meter long axon, we expect the concentration of sodium pumps in the membrane near the distal end to be roughly a 1/25 of that found near the nucleus. It is striking to note that over a length 0.1 *m*, however, the loss ratio is $e^{-.3} = .74$. This implies that material loss due to deposition and turnover becomes an issue exactly over the range of lengths of typical axons: 0.1 *m* to 1 *m*. At the lower end of this range, the axon is very homogeneous. At the higher end, one predicts significant inhomogeneity.

7 Discussion

In this paper, we used a spatial Markov chain model to unify several previous approaches to modeling fast axonal transport. After inferring the orders of magnitude of the parameters of the model, we have used standard techniques from renewal theory and linear Markov chain theory to give a precise description of the most striking qualitative feature of axonal transport: homogeneity along the length of the axon. Furthermore we predict the timescale of recovery to this homogeneous state after removing the vesicles from a segment as a function of the length scale of the segment.

In creating the model, we have become aware of two features of the biology which have not been adequately studied: 1) Deposition and decay of proteins; and 2) characterization of the process by which unbound vesicles reattach to the transport mechanism.

Regarding deposition and decay, we mentioned in Section 6 that the model unequivocally predicts a loss of material down the length of the axon. Furthermore, it seems that certain biological features that we have chosen to ignore – such as retrograde transport and a location-dependent unbinding rate [9] – would only serve to increase the non-homogeneity. In the case of sodium pumps, it is not immediately clear if there is a contradiction between deposition loss and homogeneity, but we do see a sensitive dependence on the length of the axon. Furthermore, it stands to reason that for other types of cargo the loss may be significant. We believe that this raises an interesting biological issue, and mention a few possible mechanisms for overcoming the

material loss: first, it is possible that proteins can be synthesized in axons [33]; second, there may be some signal that prevents material from depositing in a region that is already “full;” and third, it may be that for some materials “deposition” is a reversible process.

Regarding reattachment, we note that we have modeled all event wait times as 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 ATP [35] [29]. 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 [11] 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 [28] [8].

Generalizing the wait times does not significantly affect the particle perspective results of Sections 3.2 and 6.1. This is because renewal theory, which depends on independence of the wait times, is robust with respect to such changes. However, the spatial system results may be affected in that the full process is no longer Markov. While one may still expect a something like a stationary distribution for the flow-through system (where new particles enter from the nucleus and particles are removed from the system at the distal end), the analysis of the approach to equilibrium may change. In the absence of direct observation of the phenomenon, we refrain from this more detailed analysis.

In light of this known need for generalized wait times in the stepping process, it seems likely that detailed observation of the rebinding process will call for new 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.

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. This work was supported by NSF Grant DMS-061670.

References

- [1] Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2008) *The Molecular Biology of the Cell*, Garland Science, Taylor & Francis Group, New York.
- [2] Allen RD, Weiss DG, Hayden JH, Brown DT, Fijiwaki H, Simpson M (1985). 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**, 1736-1752.
- [3] Blum JJ, Reed MC (1985) A Model for Fast Axonal Transport. *Cell Motility* **5**, 507-527.
- [4] Brooks, E. A. (1999) Probabilistic methods for a linear reaction-hyperbolic system with constant coefficients. *Ann. Appl. Prob.*, **9(3)**, 719-731.
- [5] Brown A (2000) Slow axonal transport: stop and go traffic in the axon. *Nat Rev Cell Mol Biol* **1**, 153-156.
- [6] Carter NJ, Cross RA (2005) Mechanics of the kinesin step. *Nature* **435**, 308-312.
- [7] Cox DR (1962) *Renewal Theory*. Methuen & CO. Ltd., London, Section 4.5.
- [8] DeVille RL and Vanden-Eijnden E (2008) Regular Gaits and Optimal Velocities for Motor Proteins *Biophysics Journal* **95** 6: 2681-2691.
- [9] Dixit R, Ross JL, Goldman YE, Holzbaur ELF (2008) Differential regulation of dynein and kinesin motor proteins by tau. *Science* **319**, 1086-1089.
- [10] Finer JT, Simmons RM, Spudich JA (1994) Single myosin molecule mechanics: piconewton forces and nanometre steps. *Nature* **368**, 113-119.
- [11] Fisher ME and Kolomeisky (2001) Simple mechanochemistry describes the dynamics of kinesin molecules. *PNAS* **98** 14: 7748-7753.

- [12] Friedman A, Craciun G (2005) A model of intracellular transport of particles in an axon. *SIAM J Appl Math* **38**, 741-758.
- [13] Friedman A, Craciun G (2006) Approximate traveling waves in linear reaction-hyperbolic equations. *SIAM J Appl Math* **38**, 741-758.
- [14] Friedman A, Hu B (2007) Uniform convergence for approximate traveling waves in reaction-hyperbolic systems. *Indiana Math J* **56**, 2133-2158.
- [15] Friedman A, Hu B (2007) Uniform convergence for approximate traveling waves in reaction-diffusion-hyperbolic systems. *Arch Rat Mech Anal* **186**, 251-274.
- [16] Gennerich A, Carter AP, Reck-Peterson, Vale RD (2007) Force-induced bidirectional stepping of cytoplasmic dynein. *Cell* **131**, 952-965.
- [17] Gross, G.W, Weiss, D.G. (1982) "Theoretical considerations on rapid transport in low viscosity axonal regions." *Axoplasmic Transport* (Ed. D.G. Weiss), Springer-Verlag, Berlin.
- [18] Jastrow, H. 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>
- [19] Kelly, F. R. *Reversibility and stochastic networks*. John Wiley & Sons Ltd., Chichester, 1979. Wiley Series in Probability and Mathematical Statistics.
- [20] Lasek RJ, Brady ST (1982). The structural hypothesis of axonal transport: two classes of moving elements. *Axoplasmic Transport* (ed. G. Weiss), Springer-Verlag, Berlin, 1982, pp. 397-405.
- [21] Lawler G (1995) *Introduction to Stochastic Processes* Chapman and Hall, New York.
- [22] Miller R, Lasek RJ (1985) Crossbridges mediate anterograde and retrograde vesicle transport along microtubules in squid axoplasm. *J Cell Biol* **101**, 2181-2193.
- [23] Moran, D., Rowley J.C. III Visual Histology.com, Chapter 8, Nerves.
http://www.visualhistology.com/products/atlas/VHA_Chpt8_Nerves.html
- [24] Ochs S (1972) Rate of fast axoplasmic transport in mammalian nerve fibers. *J Physiol* **227**, 627-245.

- [25] Reed, M. C. and Blum, J. J. (1986) Theoretical analysis of radioactivity probes during fast axonal transport: effects of deposition and turnover. *Cell Motility and the Cytoskeleton*, **6**, 620-627.
- [26] Reed M, Blum J (1994) Mathematical Questions in Axonal Transport. In *Lectures in mathematics in the Life Sciences*, **24**, Amer. Math. Soc., Providence.
- [27] Reed, M. C., Venakides, S. and Blum, J. J. (1990) Approximate traveling waves in linear reaction-hyperbolic equations. *SIAM J. Appl. Math.*, **50**, 1671-80.
- [28] Schilstra MJ and Martin SR (2006) 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** 153-165.
- [29] Schnitzer MJ, Visscher K and Block SM (2000) Force production by single kinesin motors. *Nat. Cell. Biol.* **2** 718-722.
- [30] Stewart GH, Horwitz B, Gross GW. A chromatographic model of axoplasmic transport. In *Axoplasmic Transport* (ed. G. Weiss), Springer-Verlag, Berlin, 1982, 414-422.
- [31] Svoboda K, Schmidt C, Schnapp B (1985) Direct observation of kinesin stepping by optical trapping interferometry. *Nature* **365**, 721-727.
- [32] Takenaka T, Gotoh H (1984). Simulation of axoplasmic transport. *J Theor Biol* **107**, 579-601.
- [33] Twiss JL, Fainzilber M (2009) Ribosomes in axons - scrounging from the neighbors. *Trends in cell biology.* **19**, 236-243
- [34] Vale RD, Reese TS, Sheetz MP (1985). Identification of a novel force-generating protein, kinesin, involved in microtubule-based motility. *Cell* **42**, 39-50.
- [35] Visscher K, Schnitzer MJ, Block SM (1999) Single kinesin molecules studied with a molecular force clamp *Nature* **400**, 184-189.
- [36] Willis, WD, Grossman, RG (1981) *Medical Neurobiology: Neuroanatomical and Neurophysiological Principles Basic to Clinical Neuroscience*. Mosby, Inc.

A STOCHASTIC COMPARTMENTAL MODEL FOR FAST AXONAL TRANSPORT

LEA POPOVIC*, SCOTT A. MCKINLEY†, AND MICHAEL C. REED‡

Abstract. In this paper we develop a probabilistic micro-scale compartmental model and use it to study macro-scale 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 non-equilibrium 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?

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 [27][33], while slow transport goes at approximately 1 millimeter/day, the rate of axon growth and regeneration [6][27]. 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][29][43]. 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]. 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][12][18][40]. 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 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

*Department of Mathematics and Statistics, Concordia University, Montreal QC H3G 1M8

†Department of Mathematics, University of Florida, 358 Little Hall Box 118105, Gainesville, FL 32605 (scott.mckinley@ufl.edu)

‡Department of Mathematics, Duke University, Box 90320, Durham, NC 27701

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][34][35]. These models did not have traveling wave solutions, but the data certainly looked like approximate traveling waves. In [36] 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 co-workers have introduced new PDE models and proved these results rigorously [14][15][16][17].

Probabilistic models for axonal transport were introduced and used for simulations already in the 1980s [39][41]. However, rigorous work began with Lawler [28] 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 [31, 32] 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 [36] 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 length 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 [35]. 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 (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 [31] 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 non-equilibrium 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 δ , 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

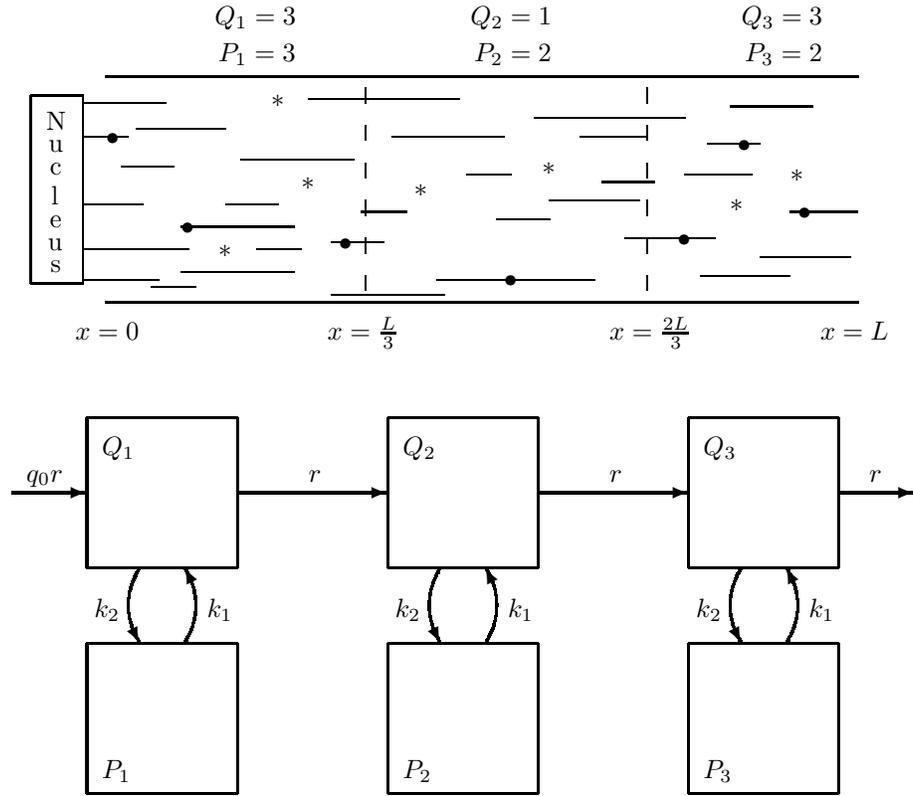
- $Q_i(t)$ is the number of particles at time t in on-transport state in section i ,
- $P_i(t)$ is the number of particles at time t in off-transport state in section i .

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 $rQ_N(t)$.

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 δ . 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 δ of each compartment. This will imply that the mean number of particles per unit length scales as ρ_0 . A graph of the model is depicted in Figure 1.

In order to insure the Markov property, we use exponential random variables for the waiting times between transition events. Specifically, we mean that after a given

FIG. 2.1. *On-and-off Transport Chain*

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. [10] 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 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 [10], 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 [10] 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 [20], Fig. 3; [21]; [30]), 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 partial differential equation. This enables us to obtain, as a special case, the asymptotic behavior of the PDE models for axonal transport in an either rate limiting or 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^δ be the number of jumps made by the particle as of time t . Observe, $X^\delta = \delta n^\delta$.

LEMMA 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

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

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

Proof. Since n^δ 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 on-transport to 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^δ 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 non-compartmental 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 2-dimensional Markov process to model the particle dynamics, where

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

DEFINITION 3.2. (Stochastic non-compartmental 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_1
- (Switch from off-transport to on-transport)
 $(X_t, 0) \rightarrow (X_t, 1)$ rate k_2
- (Lateral travel) $X_t = \int_0^t v \xi_s ds$

The path of $(X_t)_{t \geq 0}$ consists of alternating sequence 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*

$$(X_t^\delta)_{t \geq 0} \xrightarrow[\delta \rightarrow 0]{\Longrightarrow} (X_t)_{t \geq 0}$$

in distribution on the Skorokhod space of cadlag functions.

Proof. If we let ξ^δ 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 ξ^δ is a continuous-time Markov chain on $\{0, 1\}$ (independent of δ), and conditionally on ξ^δ the transition law of X is easily expressed. Likewise, for the non-compartmental model above, $(X_t, \xi_t)_{t \geq 0}$ is a strong Markov process, with ξ the same continuous-time Markov chain on $\{0, 1\}$ as ξ^δ , and conditionally on ξ the change in X is easily given in terms of its linear speed and ξ .

We will start by showing that for any $t > 0$, $(X_t^\delta, \xi_t^\delta)$ converges to (X_t, ξ_t) in distribution as $\delta \rightarrow 0$. We then show that the finite dimensional distributions of

(X^δ, ξ^δ) converge to those of (X, ξ) . A tightness argument finally implies (Lemma 16.2 and Theorem 16.3 [22]) that (X^δ, ξ^δ) converges to (X, ξ) 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 δ . 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 δ as well. This is repeated, and ξ^δ 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 τ_1 the particle behaves as if it were in active transport ($k_2 = k_1 = 0$) and conditionally on the value of τ_1 , for any $0 \leq s \leq \tau_1$ the lateral change in position over time s , $X_s^\delta - X_0^\delta$, is δn_s^δ where n_s^δ is a Poisson(rs) random variable. Moreover, by results of Proposition 3.1, conditionally on the value of τ_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 non-compartmental model. It is immediate from the definition of the model that ξ 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 ξ^δ . Since we prove our convergence in law results by first conditioning on the values of ξ^δ and ξ , we can without loss of generality henceforth assume $\xi^\delta = \xi$ a.s., and drop its superscript.

Suppose that initially the particle in the non-compartmental model is on transport at x_0 , so $(X_0, \xi_0) = (x_0, 1)$. Conditionally on τ_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 τ_1 and σ_1 the particle in both models stays in place, so conditionally on values of τ_1, σ_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 non-compartmental model, 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 η_t be the number of times the particle in either model gets back on transport until time t , and η'_t the number of times it gets off,

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

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 τ_t to t the particle is in active transport, and the same convergence argument as before implies that conditionally on the values of τ_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 η_t , $n_{\tau_t}^\delta$ is a sum of η_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 η_t i.i.d. Exponential(k_2/v) random variables. In the non-compartmental model, if $\eta'_t = \eta_t$, then conditionally on the values of η_t and τ_t , for $\tau_t \leq s \leq t$, $X_s - X_{\tau_t} = v(s - \tau_t)$ and conditionally only on the value of η_t , $X_{\tau_t}^\delta - X_0^\delta$ is a sum of η_t i.i.d. Exponential(k_2/v) variables. Hence, conditionally on η_t and τ_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 η_t , $X_{\tau_t}^\delta - X_0^\delta \Rightarrow X_{\tau_t} - X_0$.

If $\eta'_t = \eta_t + 1$, then from time τ_t to t the particle is in both models stays in place, so conditionally on η_t and τ_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 η'_t , $n_{\tau_t}^\delta$ is a sum of η'_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 η'_t i.i.d. Exponential(k_2/v) variables. In the non-compartmental model, if $\eta'_t = \eta_t + 1$, conditionally only on the value of η'_t , $X_{\tau_t}^\delta$ is a sum of η'_t i.i.d. Exponential(k_2/v) variables. Hence, conditionally only on η'_t , $X_{\tau_t}^\delta - X_0^\delta \Rightarrow X_{\tau_t} - X_0$.

Now, integrating over the possible values of η_t, η'_t and τ_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^δ, ξ^δ) and (X, ξ) , and the fact that the increments of both (X^δ, ξ^δ) and (X, ξ) are stationary.

In order to verify tightness, Theorem 16.11 [22], of the sequence of Markov processes $\{(X^\delta, \xi^\delta)\}_{\delta > 0}$, because (X^δ, ξ^δ) 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 ξ^δ 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 δ . Hence, at time h later, $\mathbb{P}\{\xi_h^\delta \neq \xi_0^\delta\} \leq 1 - e^{-kh}$. Moreover, irrespective of the value of ξ^δ , at time h later the value of $|X_h^\delta - X_0^\delta| \leq \delta n_h^\delta$ where n^δ 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 $f \in \mathcal{C}^{1,0}(\mathbb{R}_+ \times \{0, 1\})$ imply that the finite dimensional distributions of (X^δ, ξ^δ) converge to those of (X, ξ) . Verification of additional conditions, see Theorem 19.25 of [22], 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 $\mu = k_1 v / (k_2 + k_1)$, $\sigma = 2k_2 k_1 v^2 / (k_2 + k_1)^3$ 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 functional central limit theorem 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*

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

and, if B denotes a standard Brownian motion

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

in distribution on the space of continuous functions.

Proof. For these results we use the notion of stochastic averaging [23][24][26]. Note that the indicator process ξ 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, Ch 12 of [11], [19], driven by the independent indicator process ξ . The generator of (X, ξ) 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 ξ 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 $\xi = 1$: $\sigma = v, \lambda = k_2, s = 0$ and when $\xi = 0$: $\sigma = 0, \lambda = k_1, s = 1$). 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 ξ 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 ξ_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

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

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

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

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

Once we rescale the position for the particle by $1/\sqrt{n}$ and time by $1/n$, ξ still changes at a much faster rate than the position of the particle X . The generator of the rescaled centered process (X^n, ξ^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 [26]) 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 $f \in \mathcal{C}_0^2(\mathbb{R})$, $\xi_t^n \in \{0, 1\}$ 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 [26] on stochastic averaging it is only left to show 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

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

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 Δ) so that the right hand side of the inequality with $n = 1$ is less than Δ , 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 [26] 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 Partial Differential Equations Models. In order to demonstrate the connection between our model and the PDEs seen in [36][35][16], consider the process (X, ξ) 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, ξ) imply that q and p satisfy the system of PDEs

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

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

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 translation of the delta function, while the time evolution via the equations (3.1) and (3.2) will have a spreading profile. This is clear from the macroscopic limits of (X^δ, ξ^δ) 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, ξ) with a non-zero variance, as seen in Proposition 3.3 and Proposition 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 wave front 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 [36][35] that as $\varepsilon \rightarrow 0$ the solutions of

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

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

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

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

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

This asymptotic form is valid for small ε , 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 [14][15][16][17].

To see that our Proposition 3.4 provides another rigorous proof of these properties, albeit using stochastic methods, consider the process $(\varepsilon X_{\cdot/\varepsilon}, \xi_{\cdot/\varepsilon})$ (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^ε and p^ε 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 δ , 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 from (Q_1, P_1) . It follows by induction that the stationary distributions for $\{(Q_i, P_i)\}$ are independent and identically distributed 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 Sec. 8.2 of [25]. \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 non-compartmental model at any location along the axon. Namely, suppose the particles move according to the piecewise-linear Markov process (X, ξ) with a Poisson rate ρ_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 to notice 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 Δ are. Let Q_Δ and P_Δ denote the numbers of on-track and off-track particles in a segment of length Δ .

In equilibrium, Q_Δ and P_Δ are both sums of 10^ν 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_Δ and P_Δ 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 Δ is $10^\nu(\lambda_Q + \lambda_P)$. To see how homogeneous different slices of length Δ are, we consider the coefficient of variation, c_Δ , 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 [31, 32] 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 non-compartmental individual particle model with the additional feature that vesicles can move backwards as well as forwards. They compute the probability that the particle is successful and conditioning on success, the mean first hitting time. With our system-wide 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 tradeoff 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 non-empty sections of the cell at time 0. Among the particle 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, $\mathbb{E}[H_n]$ 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 non-compartmental 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 δ .

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*

$$\mathbb{E}[H'_n] = (1 - e^{-\lambda_n}) \left[\frac{1}{rq_0 p_n} + \frac{k_2}{k_1(k_2 + k_1)} \right] + \frac{1}{k_2 p_n} \left[1 - \left(\frac{r}{k_2 + r} \right)^n \left(1 + \frac{nk_2}{k_2 + r} \right) \right]$$

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 which location it was at time 0, and 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 $H'_{e,n}$ is 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 independent and identically distributed 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 δ . To this end, we fix a length ℓ and for a given compartment size ℓ , 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

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

It remains to interpret this result with respect to the parameter choices we have made. First, we must set a value for the size ℓ 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.¹

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 ℓ , the recurring ratio ℓ/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 ℓ_* 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 $1/v\rho_0 p_\ell \approx 16s$. 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 the equation (4.1) gives us is the ability 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 more 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

¹We do not wish to claim 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.

biology of deposition of materials in the cell membrane.

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 is discussed in [45], 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, $Q_i(0) + P_i(0) = 0, \forall i \in R$, is zero at time 0. Let a and ε be given small numbers and suppose that λ_∞ 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 $(Q_i, P_i), \forall i \in R$ is significantly different from λ_∞ 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 $\forall t \geq t_*$,*

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

In fact, the choice $t_ = \alpha^{-1} \ln\left(\frac{\sqrt{a|\lambda_\infty|}}{a\sqrt{\varepsilon|\lambda_\infty|-1}}\right)$ 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 ε , 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 condition for both P_i and Q_i are less than their respective equilibrium values, each are monotonically increasing functions and the above reduces to

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

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

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

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

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 $\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}_\infty + e^{-A_1 t}(\boldsymbol{\lambda}(0) - \boldsymbol{\lambda}_\infty)$ where $\boldsymbol{\lambda}_\infty = q_0rA_1^{-1}\mathbf{e}_1 = q_0(1, \frac{k_2}{k_1})$. This yields the estimate

$$|\boldsymbol{\lambda}(t) - \boldsymbol{\lambda}_\infty| \leq |e^{-A_1 t}\boldsymbol{\lambda}_\infty| \leq e^{-\alpha t}|\boldsymbol{\lambda}_\infty|$$

where α 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 Δ . 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

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

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, $\tilde{\boldsymbol{\lambda}}_\infty = q_0r\tilde{A}_1^{-1}\mathbf{e}_1 = q_0\frac{r}{\tilde{r}}(\frac{1}{k_2/k_1}) = q_010^\nu(\frac{1}{k_2/k_1})$. Both components of this vector are of order 10^ν , 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 $a^2\varepsilon|\boldsymbol{\lambda}_\infty| > 1$ requires that $|\boldsymbol{\lambda}_\infty| > 2 \times 10^3$. 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 $\ln\left(\frac{\sqrt{\varepsilon|\lambda_\infty|}}{a\sqrt{\varepsilon|\lambda_\infty|} - 1}\right)$ is contained in the interval (2.3, 2.5) and does not have a impact on how the relaxation time scales with ν . 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 ε 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 δ tends to zero, our model converges in distribution on the Skorokhod space of *càdlàg* 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 non-compartmental 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 [36, 5, 14]. 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 ATP [44] [38]. 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 [13] 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 [37] [9]. Generalizing waiting 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 Section 3.1.-3.3., nor can we use the stationary distribution employed in Section 4.1 and used for addressing the biologicals 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 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 [42]. We have also not included retrograde transport or the fact that some axons may have location-dependent unbinding rates [10]. 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. This work was supported by NSF grants DMS-061670 and EF-1038593, and an NSERC Discovery Grant.

REFERENCES

- [1] Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2008) *The Molecular Biology of the Cell*, Garland Science, Taylor & Francis Group, New York.
- [2] Allen RD, Weiss DG, Hayden JH, Brown DT, Fijiwaki H, Simpson M (1985). 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**, 1736-1752.
- [3] Blum JJ, Reed MC (1985) A Model for Fast Axonal Transport. *Cell Motility* **5**, 507-527.
- [4] Bressloff PC, (2006) Stochastic model of protein receptor trafficking prior to synaptogenesis *Physical Review E* **74** 031910.
- [5] Brooks EA (1999) Probabilistic methods for a linear reaction-hyperbolic system with constant coefficients. *Ann. Appl. Prob.*, **9**(3), 719-731.
- [6] Brown A (2000) Slow axonal transport: stop and go traffic in the axon. *Nat Rev Cell Mol Biol* **1**, 153-156.

- [7] Carter NJ, Cross RA (2005) Mechanics of the kinesin step. *Nature* **435**, 308-312.
- [8] Cox DR (1962) *Renewal Theory*. Methuen & Co. Ltd., London, Section 4.5.
- [9] DeVille RL and Vanden-Eijnden E (2008) Regular Gaits and Optimal Velocities for Motor Proteins *Biophysics Journal* **95** 6: 2681-2691.
- [10] Dixit R, Ross JL, Goldman YE, Holzbaur ELF (2008) Differential regulation of dynein and kinesin motor proteins by tau. *Science* **319**, 1086-1089.
- [11] Ethier S, Kurtz TG (1986) *Markov Processes: Characterization and Convergence*. John Wiley & Sons Inc., New Jersey
- [12] Finer JT, Simmons RM, Spudich JA (1994) Single myosin molecule mechanics: piconewton forces and nanometre steps. *Nature* **368**, 113-119.
- [13] Fisher ME, Kolomeisky (2001) Simple mechanochemistry describes the dynamics of kinesin molecules. *PNAS* **98** 14: 7748-7753.
- [14] Friedman A, Craciun G (2005) A model of intracellular transport of particles in an axon. *SIAM J Appl Math* **38**, 741-758.
- [15] Friedman A, Craciun G (2006) Approximate traveling waves in linear reaction-hyperbolic equations. *SIAM J Appl Math* **38**, 741-758.
- [16] Friedman A, Hu B (2007) Uniform convergence for approximate traveling waves in reaction-hyperbolic systems. *Indiana Math J* **56**, 2133-2158.
- [17] Friedman A, Hu B (2007) Uniform convergence for approximate traveling waves in reaction-diffusion-hyperbolic systems. *Arch Rat Mech Anal* **186**, 251-274.
- [18] Gennerich A, Carter AP, Reck-Peterson, Vale RD (2007) Force-induced bidirectional stepping of cytoplasmic dynein. *Cell* **131**, 952-965.
- [19] Griego RJ, Hersh R (1969) "Random evolutions, Markov chains, and systems of partial differential equations." *Proc. Nat. Acad. Sci. USA* **62**, 305-308.
- [20] Gross GW, Weiss, DG (1982) "Theoretical considerations on rapid transport in low viscosity axonal regions." *Axoplasmic Transport* (Ed. D.G. Weiss), Springer-Verlag, Berlin.
- [21] Jastrow H 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>
- [22] Kallenberg O (2002) "Foundations of modern probability", 2nd ed. Springer-Verlag, New York.
- [23] Khas'minskii, RZ (1966) "On stochastic processes defined by differential equations with a small parameter." *Theory Probab. Appl.* , **11**, 211-228.
- [24] Khas'minskii RZ (1966) "A limit theorem for the solutions of differential equations with random right-hand sides." *Theory Probab. Appl.* , **11**, 390-406.
- [25] Kelly FR *Reversibility and stochastic networks*. John Wiley & Sons Ltd., Chichester, 1979. Wiley Series in Probability and Mathematical Statistics.
- [26] Kurtz TG (1992) Averaging for martingale problems and stochastic approximation. *Lecture Notes in Control and Inform. Sci.* , **177**, 186-209.
- [27] Lasek RJ, Brady ST (1982). The structural hypothesis of axonal transport: two classes of moving elements. *Axoplasmic Transport* (ed. G. Weiss), Springer-Verlag, Berlin, 1982, pp. 397-405.
- [28] Lawler G (1995) *Introduction to Stochastic Processes* Chapman and Hall, New York.
- [29] Miller R, Lasek RJ (1985) Crossbridges mediate anterograde and retrograde vesicle transport along microtubules in squid axoplasm. *J Cell Biol* **101**, 2181-2193.
- [30] Moran D, Rowley JC III Visual Histology.com, Chapter 8, Nerves.
http://www.visualhistology.com/products/atlas/VHA_Chpt8_Nerves.html
- [31] Newby JM, Bressloff PC Directed Intermittent Search for Hidden Targets *New Journal of Physics* **11** 023033.
- [32] Newby JM, Bressloff PC Quasi-steady State Reduction of Molecular Motor-Based Models of Directed Intermittant Search *Bulletin of Mathematical Biology* **72** 1840-1866.
- [33] Ochs S (1972) Rate of fast axoplasmic transport in mammalian nerve fibers. *J Physiol* **227**, 627-245.
- [34] Reed MC, Blum JJ (1986) Theoretical analysis of radioactivity probes during fast axonal transport: effects of deposition and turnover. *Cell Motility and the Cytoskeleton*, **6**, 620627.
- [35] Reed MC, Blum J (1994) Mathematical Questions in Axonal Transport. In *Lectures in mathematics in the Life Sciences*, **24**, Amer. Math. Soc., Providence.
- [36] Reed MC, Venakides S, Blum JJ (1990) Approximate traveling waves in linear reaction-hyperbolic equations. *SIAM J. Appl. Math.*, **50**, 167180.
- [37] Schilstra MJ, Martin SR (2006) 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** 153-165.
- [38] Schnitzer MJ, Visscher K, Block SM (2000) Force production by single kinesin motors. *Nat.*

- Cell. Biol.* **2** 718-722.
- [39] Stewart GH, Horwitz B, Gross GW (1982) A chromatographic model of axoplasmic transport. In *Axoplasmic Transport* (ed. G. Weiss), Springer-Verlag, Berlin, 1982, 414-422.
 - [40] Svoboda K, Schmidt C, Schnapp B (1985) Direct observation of kinesin stepping by optical trapping interferometry. *Nature* **365**, 721-727.
 - [41] Takenaka T, Gotoh H (1984) Simulation of axoplasmic transport. *J Theor Biol* **107**, 579-601.
 - [42] Twiss JL, Fainzilber M (2009) Ribosomes in axons - scrounging from the neighbors. *Trends in cell biology*. **19**, 236-243
 - [43] Vale RD, Reese TS, Sheetz MP (1985). Identification of a novel force-generating protein, kinesin, involved in microtubule-based motility. *Cell* **42**, 39-50.
 - [44] Visscher K, Schnitzer MJ, Block SM (1999) Single kinesin molecules studied with a molecular force clamp *Nature* **400**, 184-189.
 - [45] Welte MA, Gross SP (2008) Molecular motors: a traffic cop within? *HFSP Journal* **2** 4: 178-182.
 - [46] Willis WD, Grossman RG (1981) *Medical Neurobiology: Neuroanatomical and Neurophysiological Principles Basic to Clinical Neuroscience*. Mosby, Inc.