PROCEEDINGS

Open Access

Analytical results on the Beauchemin model of lymphocyte migration

Johannes Textor[1,2*], Mathieu Sinn[3], Rob J de Boer[1]

Abstract

The Beauchemin model is a simple particle-based description of stochastic lymphocyte migration in tissue, which has been successfully applied to studying immunological questions. In addition to being easy to implement, the model is also to a large extent mathematically tractable. This article provides a comprehensive overview of both existing and new analytical results on the Beauchemin model within a common mathematical framework. Specifically, we derive the motility coefficient, the mean square displacement, and the confinement ratio, and discuss four different methods for simulating biased migration of pre-defined speed. The results provide new insight into published studies and a reference point for future research based on this simple and popular lymphocyte migration model.

Background

A unique property of the immune system is that it mainly consists of constantly moving cells. Technological progress has facilitated the study of lymphocyte migration at increasingly higher resolutions, culminating in the application of two-photon microscopy to tracking single lymphocytes in the living animal [1-3]. Accompanying this progress, mathematical and computational models of lymphocyte migration have been developed and applied to both qualitative and quantitative questions. These models can be roughly categorized as follows according to their level of detail: Whole-population models, usually formulated using ordinary or partial differential equations, do not distinguish between individual cells, but treat cell subpopulations as continuous quantities. Particle-based models do consider each cell individually, but treat cells as freely moving particles without mass, volume, and shape. Finally, whole-cell models such as the Cellular Potts Model [4] explicitly represent the cell and its interaction with the environment. In the present paper, we analyze a particle-based model that was introduced by Beauchemin,

Dixit, and Perelson [5]. Throughout the paper, we refer to this model as the "Beauchemin Model".

Particle-based approaches such as the Beauchemin model are used for studying questions where a whole-population approach does not provide sufficient information, but a whole-cell model is not necessary, not feasible for computational reasons or would require too many assumptions on unknown parameters. For instance, Grigorova et al. [6] and ourselves [7] recently used the Beauchemin model to study the transit of T cells through a lymph node in the absence of antigen, and we used it to determine the amount of directional bias that could be detected in random T cell migration using contemporary two-photon imaging techniques [7].

The Beauchemin model is defined by three parameters - a speed v_{free}, a time interval t_{free}, and another time interval t_{pause}. The model describes a particle moving freely in a three-dimensional space according to the following rules: The particle starts at a fixed position and turns in a random direction (more precisely, a direction is sampled from a uniform distribution on the unit sphere). The turn takes a fixed time t_{pause}, during which the cell does not move. This reflects the time it takes to "displace the internal structure which brings about motion" [4]. Then, the cell moves in the chosen direction with velocity v_{free} for a fixed time t_{free}, after which it stops and restarts the process.

* Correspondence: johannes.textor@gmx.de
[1]Theoretical Biology & Bioinformatics, Utrecht University, 3508 TC Utrecht, The Netherlands
Full list of author information is available at the end of the article

The main distinguishing characteristic of the model is the introduction of the pause phase. In this sense, it can be seen as a generalization of the "ideal chain" model of polymer physics [8], in which a chain of rigid rods of fixed length ℓ is considered that are freely jointed to each other. This is mathematically equivalent to the Beauchemin model if we set $v_{\text{free}} \cdot t\text{free} = \ell$ and $t_{\text{pause}} = 0$. In particular, analytical results e.g. on the mean square displacement of the ideal chain [8] are directly transferable to this case. As we shall see later on in this paper, for some quantitative properties even the more general case with $t_{\text{pause}} \geq 0$ can be reduced to the ideal chain model.

This paper is structured as follows: In the upcoming section, we begin by formulating a mathematical framework for our analysis, which explicates sufficiently but not overly general notions of random walks and Brownian motion. We deliberately include a fairly large amount of detail, including proofs, in order to present our results in a self-contained fashion. Thereafter, we present analytical results about random motion in the Beauchmin model, and discuss four different ways for simulating biased migration with pre-defined speed. We wrap up with concluding remarks.

We point out that around half of the results discussed in this paper originate from our recent study that used the Beauchemin model [7]. We chose to integrate these results into the present paper in rewritten and extended form because they were formerly only published in condensed form as supporting online material. In this manner, we hope to provide a coherent account of all available analytical results, which we anticipate to serve as a useful reference for future users of the model. We mark results that have appeared earlier in our previous study or elsewhere (e.g. Proposition 10) by giving the appropriate reference in proposition or theorem statements, while propositions or theorems stated without such a reference (e.g. Proposition 12) have, to our knowledge, not been published before.

Mathematical foundations

In this section, we outline the basic mathematical framework that we use to describe stochastic cell motion and, later on, the Beauchemin model. A similar, but slightly less general framework (based on random walk on a lattice) is used in Berg's textbook *Random Walks in Biology* [9], while a much more general account of random walks than given here is usually presented in probability textbooks (e.g., [10]). At the end of this section, we will outline the connection of this random walk to the corresponding macroscopic partial differential equation model, which is the convection-diffusion-equation.

First of all, we introduce some notational conventions. We denote the expectation of a random variable ξ by E

$[\xi]$, its variance by $\text{Var}(\xi) = \text{E}\left[(\xi - \text{E}[\xi])^2\right]$ and its probability density function by f_ξ. Elementary random variables are written in lowercase Greek letters (usually ξ_i) while derived random variables that are functions of elementary random variables are written in capital Latin letters. All random variables are assumed to be continuous and real-valued. Mean and variance of random variables will also be written as μ and σ^2, if the referenced variable is clear from context.

Observed tracks and associated measures

The Beauchemin model is usually validated against, or used to generate, cell tracking data as it would be recorded in a two-photon microscopy experiment. As we will see later, there can be subtle differences between the real underlying state and the observed state of a cell. We thus make this difference explicit in our analysis by means of the following definitions for cell tracking terminology.

Definition 1 (Cell track). *A d-dimensional cell track is a finite sequence*

$$(x(1), t(1)), \ldots, (x(k), t(k))$$

consisting of positions $x(1), \ldots, x(k) \in \mathbb{R}^d$ *and increasing time points* $t(1), \ldots, t(k) \in \mathbb{R}$.

The time indices are written in brackets to indicate that the positions are discrete-time observations of an underlying continuous motion process. For population-based cell track analysis, it is common to align a set of tracks to a common starting point.

Definition 2 (Aligned cell track sets). *For a cell track* $T = (x(1), t(1)), \ldots, (x(k), t(k))$, *the track that results from subtracting the first element from all elements in the sequence, i.e.*

$$T_0 = (0, 0), (x(2) - x(1), t(2) - t(1)), \ldots, (x(k) - x(1), t(k) - t(1))$$

is called the zero-aligned version of T.

Throughout the paper, we will assume that all cell track sets are zero-aligned.

Definition 3 (Mean displacement and mean square displacement). *Let* $x_1(t), \ldots, x_n(t)$ *be the positions of* n *zero-aligned cell tracks at some fixed time* t. *Then the mean displacement* $D(t)$ *for S is defined by*

$$D(t) = \frac{||x_1(t)|| + \cdots + ||x_n(t)||}{n}$$

where $|| \cdot ||$ *is the d-dimensional euclidian norm or vector length, and the mean square displacement* $D^2(t)$ *is defined by*

$$D^2(t) = \frac{||x_1(t)||^2 + \cdots + ||x_n(t)||^2}{n}$$

Beauchemin et al. [5] estimated the three parameters v_{free}, t_{free} and t_{pause} of their model by fitting it to mean displacement data: For each combination of parameters, 10^6 simulations of the model were run, and the mean displacement data was recorded and averaged over all simulations. The resulting mean displacement curve was compared to real data from several publications. The parameter triplets were ranked according to their squared sum of errors to the real data (see Table 1). It was found that the values t_{free} = 2.0 min, t_{pause} = 0.5 min and v_{free} = 18.8 μm/min gave the best agreement to the real data. These values are biologically reasonable and similar to those estimated by Miller et al. [1].

Microscopic random walk with bias

In this section, we will define the notion of a one-dimensional random walk used in this paper and derive its relation to a one-dimensional convection-diffusion process. These concepts can be easily generalized to higher dimensions under the assumption that the dimensions are pairwise uncorrelated, while stochastic independence in the strict sense is not necessary. This assumption holds for the particular random walks we analyze in this paper.

Definition 4 (One-dimensional random walk [9,11]). *Let ξ_i, $i \in \mathbb{N}$, be independent, identically distributed real-valued random variables with mean μ and variance σ^2. Then the sequence (S_t), $i \in \mathbb{N}$, defined by*

$$S_t = \sum_{i=1}^{t} \xi_i$$

is called a random walk with step bias μ and step variance σ^2. If μ =0, (S_t) is called an unbiased random walk.

Intuitively, the sequence (S_t) is the random walk trajectory and the ξ_i are the individual steps. The above definition is a slightly generalized version of the commonly used entirely discrete random walk with fixed steps to the left and the right: We assume that each

step is sampled from a (possibly continuous) random variable with finite first and second moments. The following corollary provides an important characteristic of this random walk:

Corollary 5 (Mean and variance of a random walk). *For a random walk (S_t) as given by Definition 4,*

$$E[S_t] = t\mu \ and \ Var(S_t) = t\sigma^2.$$

Proof. Since the ξ_i are independent and thus uncorrelated, this is a direct consequence of the linearity of variance and expectation. □

If the random walk is unbiased, $Var(S_t)$ can be geometrically interpreted as the expected mean square squared displacement of the particles at time t (Definition 3). This leads to an important characteristic of the unbiased random walk: In a particle ensemble, the particles will go nowhere on average, and their expected mean square displacement is linear in time. This characterization is often used in the immunological literature [2] to convey the difference between a random walk versus directed migration, where the *root* mean square displacement is linear in time (Figure 1).

For sufficiently long observations, we do not only know the variance and mean of a random walk, but can furthermore obtain a good approximation of the entire particle distribution from the central limit theorem.

Theorem 6 (Central limit theorem [9,10]). *Let (S_t) be a random walk as given in Definition 4. Then as $t \to \infty$, the distribution of the random variable*

$$Z_t = \frac{S_t - t\mu}{\sqrt{t}}$$

converges to a normal distribution with mean 0 and variance σ^2.

Thus, after a sufficiently long time, the distribution of the random walk (S_t) at fixed times is approximately normal. Of course, the terms "sufficiently long" and "approximately normal" are rather vague. In the context of intravital two-photon imaging, which is limited to a finite, often axially thin imaging region, we can usually not safely assume that we are in the range of validity of the normal distribution approximation: We are observing only a few "steps" (prolonged periods of rather persistent motion followed by short pauses) of a random walk. Beltman et al. [3] and ourselves [7] have discussed the implications of this technical limitation for detecting directional bias and estimating the motility coefficient.
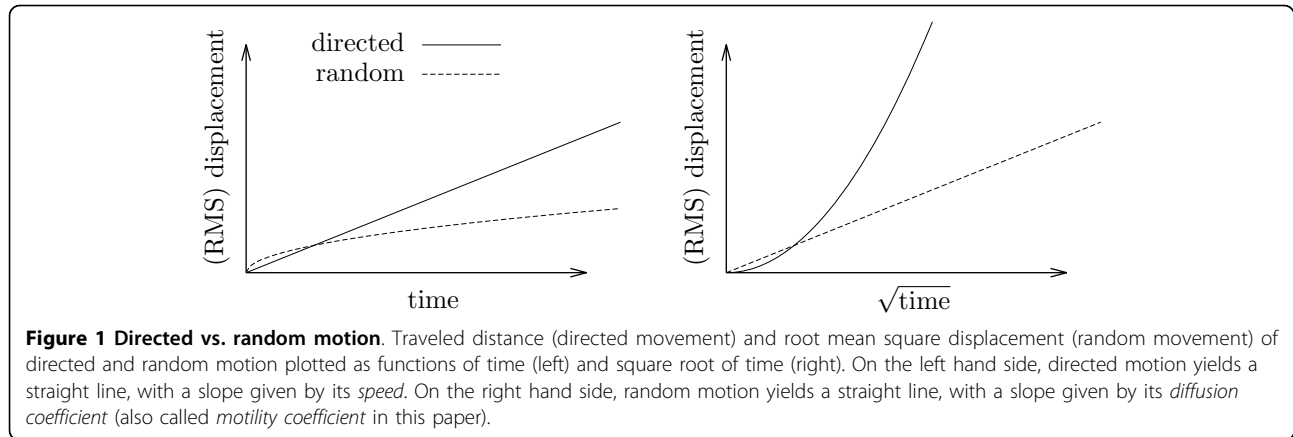
We now generalize the random walk to higher dimensions.

Definition 7 (Multi-dimensional random walk [7]). *Let $d \in \mathbb{N}$. Suppose $(\xi_i^{(1)}, \ldots, \xi_i^{(d)})$, $i \in \mathbb{N}$ are independent, identically distributed $\mathbb{R}^d$-valued random vectors*

Table 1 Different parameters can lead to similar motility coefficients.

Rank	t_{pause} min	t_{free} min	v_{free} μm/min	M μm²/min	SSR
1st	0.5	2.0	18.8	94.25	1.00
2nd	0.5	2.5	16.6	95.68	1.01
3rd	0.25	2.0	17.6	91.78	1.02
4th	1.5	1.5	26.1	92.89	1.03
5th	0.75	1.5	23.8	94.41	1.03

Parameter triplets of the 3D cell random walk model that were found to best fit *in vivo* data by Beauchemin et al. [5] along with their corresponding motility coefficients as per equation (2). Beauchemin et al. calculated sums of squared residuals (SSR) to a set of *in vivo* mean displacement data from four publications [24,25,1,26]. The SSR values shown here are normalized to the best ranking triplet.

Figure 1 Directed vs. random motion. Traveled distance (directed movement) and root mean square displacement (random movement) of directed and random motion plotted as functions of time (left) and square root of time (right). On the left hand side, directed motion yields a straight line, with a slope given by its *speed*. On the right hand side, random motion yields a straight line, with a slope given by its *diffusion coefficient* (also called *motility coefficient* in this paper).

with $\xi_i^{(1)}, \ldots, \xi_i^{(d)}$ pairwise uncorrelated, $E[\xi_i^{(1)}] = \mu^{(1)}, \ldots, E[\xi_i^{(d)}] = \mu^{(d)}$ and $Var(\xi_i^{(1)}) = \ldots = Var(\xi_i^{(d)}) = \sigma^2$. Then the sequence (S_t), $t \in \mathbb{N}$, defined by

$$S_t = (S_t^{(1)}, \ldots, S_t^{(d)}) = \sum_{i=1}^{t} (\xi_i^{(1)}, \ldots, \xi_i^{(d)})$$

is called a d-dimensional random walk with step bias $(\mu^{(1)}, ..., \mu^{(d)})$ and step variance σ^2. If $\mu^{(1)} = ... = \mu^{(d)} = 0$, then (S_t) is called an unbiased d-dimensional random walk.

The following theorem gives the limit distribution of suitably scaled multi-dimensional random walks within our framework.

Theorem 8 (Multi-dimensional central limit theorem [7]). *Let (S_t) be a d-dimensional random walk as given in Definition 7. Then as $t \to \infty$, the distribution of the random vector*

$$Z_t = \frac{(S_t^{(1)} - t\mu^{(1)}, \ldots, S_t^{(d)} - t\mu^{(d)})}{\sqrt{t}}$$

converges to a d-dimensional normal distribution with zero means and covariance matrix

$$\Sigma = \sigma^2 E$$

Proof. Let $a_1, \ldots, a_d \in \mathbb{R}$. According to the Cramér-Wold Theorem, it suffices to show that $(a_1(S_t^{(1)} - t\mu^{(1)}) + \ldots + a_d(S_t^{(d)} - t\mu^{(d)}))/\sqrt{t}$ converges in distribution to a normal random variable with zero mean and variance $(a_1^2 + \ldots + a_d^2)\sigma^2$ (see [12]). Note that the random variables $a_1\xi_i^{(1)} + \ldots + a_d\xi_i^{(d)}$, $i \in \mathbb{N}$, are independent, identically distributed and have mean $a_1\mu^{(1)} + ... + a_d\mu^{(d)}$. Furthermore, since $\xi_i^{(1)}, \ldots, \xi_i^{(d)}$ are pairwise uncorrelated,

$$Var(a_1\xi_i^{(1)}, \ldots, a_d\xi_i^{(d)}) = (a_1^2 + \ldots + a_d^2)\sigma^2.$$

Hence, the result follows from the central limit theorem for one-dimensional random walks. □

The central limit argument shows that even complex cell migration processes might have a simple description when observed for a sufficiently long time: The diffusion coefficient is a function of all "microscopic" cell motility parameters such as persistence time, speed, turning angle distributions and so on. In the scaling limit, this single coefficient describes an unbiased random walk completely. Therefore, it is more common to describe random walk processes on such timescales in terms of a differential equation instead of a stochastic process. This can be done using a *scale transition*, as we will explain below.

The convection-diffusion equation

The random walk (S_t) introduced in Definition 4 is a plausible microscopic model for the position of particles. However, if the positions are observed on large time scales (e.g., only every 20th state of the random walk can be observed due to a limited temporal resolution), then a macroscopic continuous model can be used to describe the migration. This continuous model no longer depends on the detailed characteristics of the microscopic model such as the turning angle distribution or persistence length (i.e., the precise shape of the ξ_i).

Mathematically, we construct the continuous model as follows: For $n \in \mathbb{N}$ let $(Z_n(t))$ with $t \geq 0$ be given by

$$Z_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor tn \rfloor} (\xi_i - \mu) + t\mu. \qquad (1)$$

We write the time index t in brackets in order to emphasize that the processes are given in continuous time. The process $(Z_n(t))$ is obtained by observing n steps of the random walk (S_t) in every time interval $(t - 1, t]$ with $t \in \mathbb{N}$. It is scaled such that $E[Z_n(t)] = E[S_t] = t\mu$

and $\text{Var}(Z_n(t)) = \text{Var}(S_t) = t\sigma^2$ for every $t \in \mathbb{N}$. Between two successive steps, the process is constant.

Let $(Z(t))$, $t \geq 0$, denote the limiting process obtained as n, the number of steps per time interval, tends to infinity. According to the Functional central limit theorem (sometimes also referred to as *Donsker's Theorem*, see [12]), $(Z(t))$ has the following two properties:

- Any finite-dimensional distribution of $(Z(t))$ is Gaussian.
- $\text{E}[Z(t)] = t\mu$ and $\text{Cov}(Z(s), Z(t)) = \min(s, t)\sigma^2$ for $s, t \geq 0$.

A process with these properties is called *Brownian motion* with *drift* μ. If $t>0$, then $Z(t)$ has a probability density $\varphi(x, t)$ given by

$$\phi(x,t) = \frac{1}{\sqrt{2\pi t}\sigma}\ \exp\left(-\frac{(x - t\mu)^2}{2t\sigma^2}\right)$$

for $x \in \mathbb{R}$.

As one can easily verify, this probability density function solves the following well-known differential equation for the initial condition $\varphi(x, 0) = \delta(x)$, where δ denotes the Dirac delta distribution.

Theorem 9 (Convection-diffusion equation [11,13]). *Let $C = \mu$ and $M = \sigma^2/2$. Then*

$$\frac{\partial}{\partial t}\phi(x,t) = \left[-C\frac{\partial}{\partial x}\ + M\frac{\partial^2}{\partial x^2}\right]\ \phi(x,t).$$

Taken the initial condition given above, the solution of the convection-diffusion-equation exists and is unique [13].

As the sum of independent Brownian motions is again a Brownian motion, the convection-diffusion equation describes the dynamics of single diffusing particles as well as of ensembles of particles carrying out motions independently: Assuming the ensemble is large enough such that stochastic perturbations can be ignored, the convection-diffusion-equation describes the evolution of the particle concentration over time. This scaling argument justifies the use of the convection-diffusion equation for simulating large cell populations on timescales that are substantially longer than the rhythmicity of the cell migration process. For example, we used the convection-diffusion equation to simulate the transit of T cells through the lymph node, which takes approximately half a day [7].

Note that in the context of molecular motion, the quantity M is usually called *diffusion coefficient*. Throughout this paper we will use the term *motility coefficient* to indicate that we are talking about particles that are substantially larger than molecules; this is

common in the immunological literature since Miller et al.'s early work [1].

Similarly as for the discrete model, the continuous model can be generalized to multiple dimensions. In this case, $(Z_n(t))$ is obtained by observing n steps of a d-dimensional random walk (S_t) in every time interval of unit length. For n tending to infinity, the Functional central limit theorem yields convergence to a d-dimensional Brownian motion with independent components and convection coefficients $C^{(1)} = \mu^{(1)}$, ..., $C^{(d)} = \mu^{(d)}$. The motility coefficient is equal to $M = d\sigma^2/2$, where d is the number of observed dimensions.

Summary
As a microscopic model for stochastic motility of particles, we have first introduced one-dimensional random walks. According to this model, particles move by taking independent and identically distributed steps at equidistant discrete time points. Therefore, just like in the Beauchemin model, our framework does not consider persistence to last longer than the duration of a single step; the model is fully characterized by the absolute time between successive steps and the distribution of steps. Important characteristics of the step distribution are the step bias μ and the step variance σ^2. The random walk is called unbiased if $\mu = 0$, and biased, otherwise. If we observe many particles carrying out simultaneous unbiased random walks, the particles will go nowhere on the average, but they will spread out increasingly, their distribution being approximately normal. The model can be easily generalized to d dimensions if we assume that the steps are dimension-wise uncorrelated (not necessarily statistically independent). Essentially, we can then treat the d dimensions like d separate one-dimensional random walks.

On the macroscopic scale, (biased) random walks can be approximated by Brownian motion (with drift). In this case the motion of particles is described sufficiently well by only two parameters, namely the *motility coefficient* $M = \sigma^2/2$ and the *convection coefficient* $C = \mu$. The motility coefficient M quantifies the random component of a biased random walk, while the convection coefficient C (which can be interpreted as a velocity) characterizes the directed component, if there is any (Figure 2). Convergence to Brownian motion can also be obtained for much more general random walk models using appropriate central limit theorems. However, the particular random walk chosen here is general enough for our purpose of analyzing the Beauchemin model.

The Beauchemin model
In this section, we analytically derive three quantitative characteristics of the particular random walk defined by the Beauchemin model, namely the motility coefficient,

the mean square displacement, and the confinement ratio.

Motility coefficient

The fact that the model is three-dimensional might seem to complicate our analysis. However, fortunately, it is possible to decouple the dimensions of the model, as will be shown in the following. The upcoming result was stated by ourselves [7] and, more recently, by Donovan and Lythe [14].

Proposition 10 (Motility Coefficient of the Beauchemin model [7,14]). *Let* $(\xi_i, v_i, \zeta_i) \in \mathbb{R}^3$, $i \in \mathbb{N}$, *be independent random vectors distributed uniformly on a sphere with radius r. Set* $(X_t, Y_t, Z_t) = \sum_{i=1}^{t} (\xi_i, v_i, \zeta_i)$. *Then* X_t, Y_t *and* Z_t *are unbiased random walks as defined in 4, and as* $t \to \infty$, *the distribution of* $(X_t, Y_t, Z_t)/\sqrt{t}$ *converges to a 3-variate normal distribution with zero mean and covariance matrix*

$$\sum = \begin{pmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix}, \ \sigma^2 = r^2/3.$$

Proof. The key ingredients to the proof are

(i) ξ_i, v_i, ζ_i are uniformly distributed on $[-r, r]$.
(ii) $\mathrm{Cov}(\xi_i, v_i) = \mathrm{Cov}(\xi_i, \zeta_i) = \mathrm{Cov}(v_i, \zeta_i) = 0$.

In order to establish (i), note that $P(\xi_i < h)$, i.e., the probability that ξ_i lies below the height h in the sphere, is proportional to the part of the sphere's surface area lying below the plane $\xi_i = h$. The surface of a spherical cap is proportional to its height, and thus ξ_i must be uniformly distributed on $[-r, r]$ (see [15]). By isotropy, the same arguments hold for v_i and ζ_i. The statement (ii) on pairwise uncorrelatedness is obtained by (ii) is obtained by observing that (ξ_i, v_i) is uniformly distributed on a circle with radius $r - |z|$ given that $\zeta_i = z$, and hence $\mathrm{Cov}(\xi_i, v_i) = 0$.

Now, (i) implies $\mathrm{E}[\xi_i] = \mathrm{E}[v_i] = \mathrm{E}[\zeta_i] = 0$, and hence (X_t), (Y_t) and (Z_t) are unbiased random walks. The statement on the asymptotic distribution of (X_t, Y_t, Z_t) follows by (ii) together with Theorem 8, and the fact that a random variable with uniform distribution on $[-r, r]$ has variance $r^2/3$. □

Hence, particles in the Beauchemin model behave exactly like the superposition of three uncorrelated random walks with uniformly distributed steps.

Mean square displacement

Putting together our result from the last section with the relationship between the motility coefficient M and the dimension-wise step variance σ^2 of a random walk, we obtain the following relation between the diffusion coefficient and the microscopic parameters:

$$M = \frac{(v_{\text{free}} \times t_{\text{free}})^2}{6(t_{\text{free}} + t_{\text{pause}})} \tag{2}$$

The corresponding diffusion coefficients for the 5 best-fitting parameter triplets determined by Beauchemin et al. are given in Table 1. It turns out that all the best-fitting triplets have very similar motility coefficients. On the other hand, the microscopic motility parameters vary quite a lot: For example, the pause time of the 4th best fitting triplet is 5 times longer than that of the 3rd best fitting triplet. The triplets were ranked by Beauchemin et al. according to the sum of squared residuals (SSR) to the experimental data. However, the SSR of the 5th best fitting triplet is only less than 4% higher than that of the best fitting triplet. Our upcoming analysis will provide some insight why this occurs.

The next two propositions will yield a precise characterization of the *variance* - and with it the mean square displacement - of the position of single particles and ensembles of observed particles in the Beauchemin model. It is mathematically more convenient to analyze the mean square displacement rather than the mean displacement because, as we have shown above, the
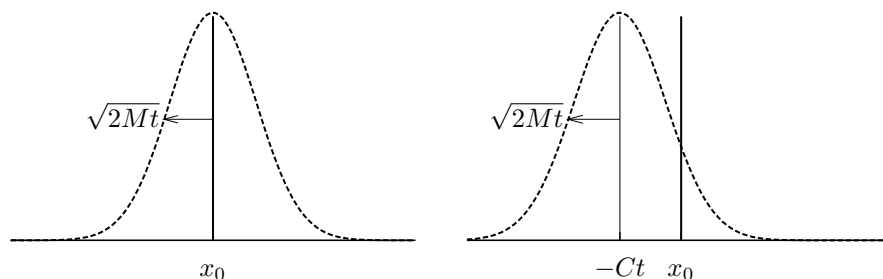

Figure 2 Position distributions for biased random walks. Expected distribution of zero-aligned cell displacements measured in one dimension in the case of an unbiased random walk with motility coefficient M (left) and a biased random walk with convection coefficient C (right) as per theorem 9. In both cases, the distribution is Gaussian with standard deviation $\sqrt{2tM}$. In the biased case, the mean of the distribution moves away from the origin with time.

dimensions can be decoupled. For mean displacement, this does not hold, and we would have to take into account the stochastical dependence between dimensions.

We start out with the variance of the position of one particle that starts its random walk at time $t = 0$. For simplicity, we will derive the following equations assuming that one dimension is observed. Due to the independence of the dimensions discussed above, multi-dimensional versions of these results are obtained simply by multiplying the derived quantity (i.e., mean square displacement or confinement ratio) with the number of observed dimensions.

Proposition 11 (Expected square displacement of single particles in the Beauchemin model). *A particle starts at time $t = 0$ in position $x(0) = 0$ to carry out a random walk as defined by the Beauchemin model with fixed parameters $t_{pause}, t_{free}, v_{free}$. Let $x(t + \tau)$ denote the position of the particle observed in one dimension at time $t + \tau$ where t is an integer multiple of $t_{free} + t_{pause}$ and $0 \leq \tau < t_{free} + t_{pause}$. Then the expected square displacement of the particle is given by:*

$$E\left[D^2(t + \tau)\right] = Var(x(t + \tau)) = 2Mt + \frac{v_{free}^2}{3}(\max(\tau - t_{pause}, 0))^2$$

with M as defined in equation (2).

Proof. The term $2Mt$ follows directly from Proposition 10 and the linearity of variance (see also Corollary 5). Thus we continue with $t = 0$. Then we have from Proposition 10 that the position at time $t_{free} + t_{pause}$ is

uniformly distributed on the interval $[-v_{free}t_{free}, v_{free}t_{free}]$, having variance $2M(t_{free} + t_{pause})$.

For $\tau \leq t_{pause}$, the particle's position does not change, hence its variance is equal to 0. For $\tau > t_{pause}$, the position is uniformly distributed on an interval with a size proportional to $\tau - t_{pause}$. Hence the variance has quadratic form $y = a(\tau - t_{pause})^2$. Inserting $\tau = t_{free} + t_{pause}$, we obtain

$$2M(t_{free} + t_{pause}) = a\,t_{free}^2$$

Inserting the definition of M from above, this results in

$$a = \frac{(v_{free})^2}{3} \qquad (3)$$

Combining the cases $\tau \leq t_{pause}$ and $\tau > t_{pause}$, we arrive at the proposition. □

A plot of the expected square displacement shows a series of iterated pulses that approaches the linear progression described by the diffusion equation from below (see Figure 3). However, we will most likely not see such phasic behaviour if we take the average of several observed particles, since the particles will in general not be synchronized and the pulses that are caused by the free runs between turns will average out. Consequently, the next step in our analysis is to derive the expected mean square displacement for a particle ensemble. We do this by assigning to each observed particle a *phase*, which reflects the particle's state at the time when we
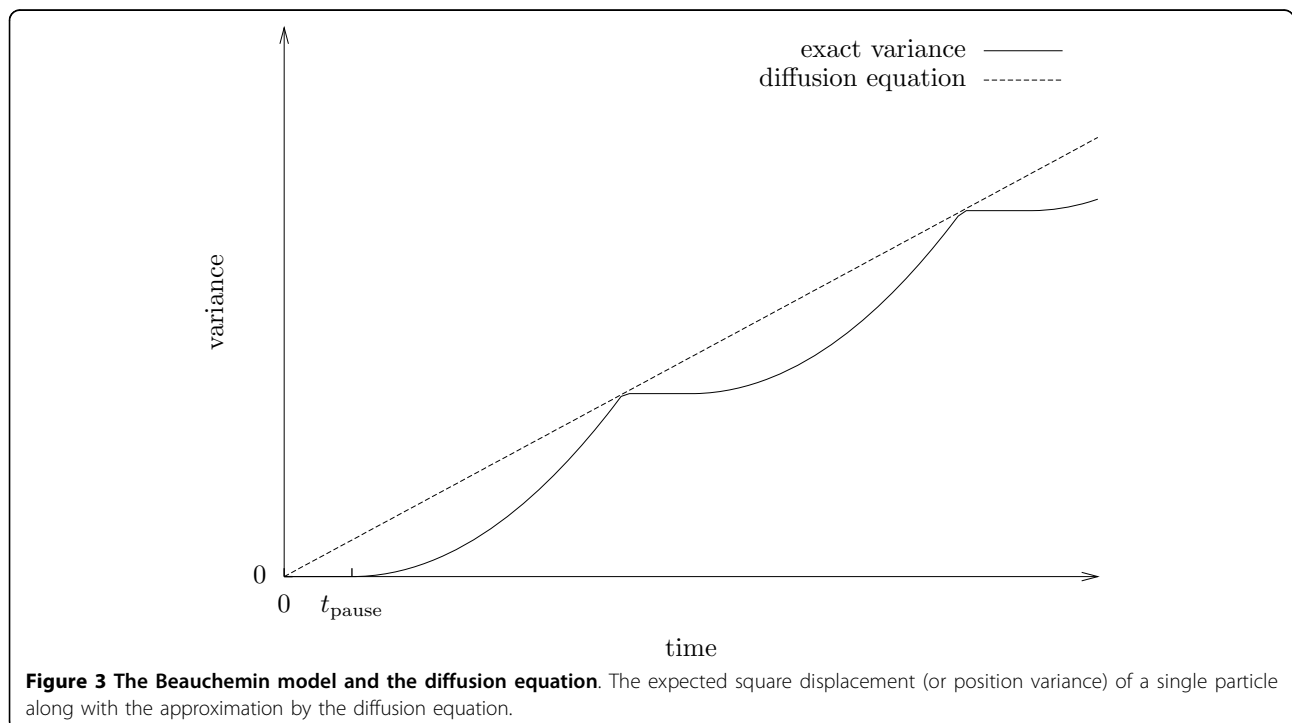

Figure 3 The Beauchemin model and the diffusion equation. The expected square displacement (or position variance) of a single particle along with the approximation by the diffusion equation.

start our observation. Then we average over all possible phases.

Proposition 12 (Empirical mean square displacement of the Beauchemin model). *Set $t_{run} = t_{free} + t_{pause}$. A particle starts at a uniformly distributed random time $-t_0$, $-t_{run} < -t_0 \leq 0$ in some position to carry out a random walk as defined by the Beauchemin model with fixed parameters t_{pause}, t_{free}, v_{free}. Let $x(t) \in \mathbb{R}$ denote the position of this particle observed in one dimension at time t, where we choose the coordinate system such that $x(0) = 0$. Then the expected square displacement of the particle at time t is given by*

$$E[D^2(t)] = \frac{1}{t_{run}} \int_{-t_{run}}^{0} Var(x(t-t_0) - x(t_0)) \, dt_0$$

$$= 2Mt - 2Mt_{free} \times \begin{cases} \frac{1}{3} & t \geq t_{free} \\ \frac{1}{3}\left(\frac{t}{t_{free}}\right)^3 - \left(\frac{t}{t_{free}}\right)^2 + \left(\frac{t}{t_{free}}\right) & t < t_{free} \end{cases}.$$

Proof. See Figure 4 for an illustration of how the parameter δ affects the variance of the observed particle position during the initial observation time. For $\delta \leq 0$, the cell is resting at the beginning of observation time and remains paused for time $t_{pause} - \delta$. For $\delta > 0$, the cell is moving at the beginning of observation time and keeps moving for time δ. Let us write $Var(S_\delta(t))$ for the position variance of a particle with phase shift δ.

Denote by τ_0 the time that the particle was already moving when we started observing it. Hence,

$$\hat{x}(t) = x(t + \tau_0) - x(\tau_0)$$

where $\hat{x}(t)$ denotes our observed position. We are interested in the variance $Var[\hat{x}(t)]$, for which the following holds:

$$Var[\hat{x}(t)] \tag{4}$$

$$= Var[x(t + \tau_0) - x(\tau_0)] \tag{5}$$

$$= Var[x(t + \tau_0)] + Var[x(\tau_0)] - 2Cov[x(t + \tau_0), x(\tau_0)] \tag{6}$$

What we are deriving is the expectation of expression 6, which is a function of the random variable τ_0. Due to the markov property of the random walk, only the last observed step matters and we can take τ_0 as uniformly distributed over $[0, t_{free} + t_{pause})$. Thus,
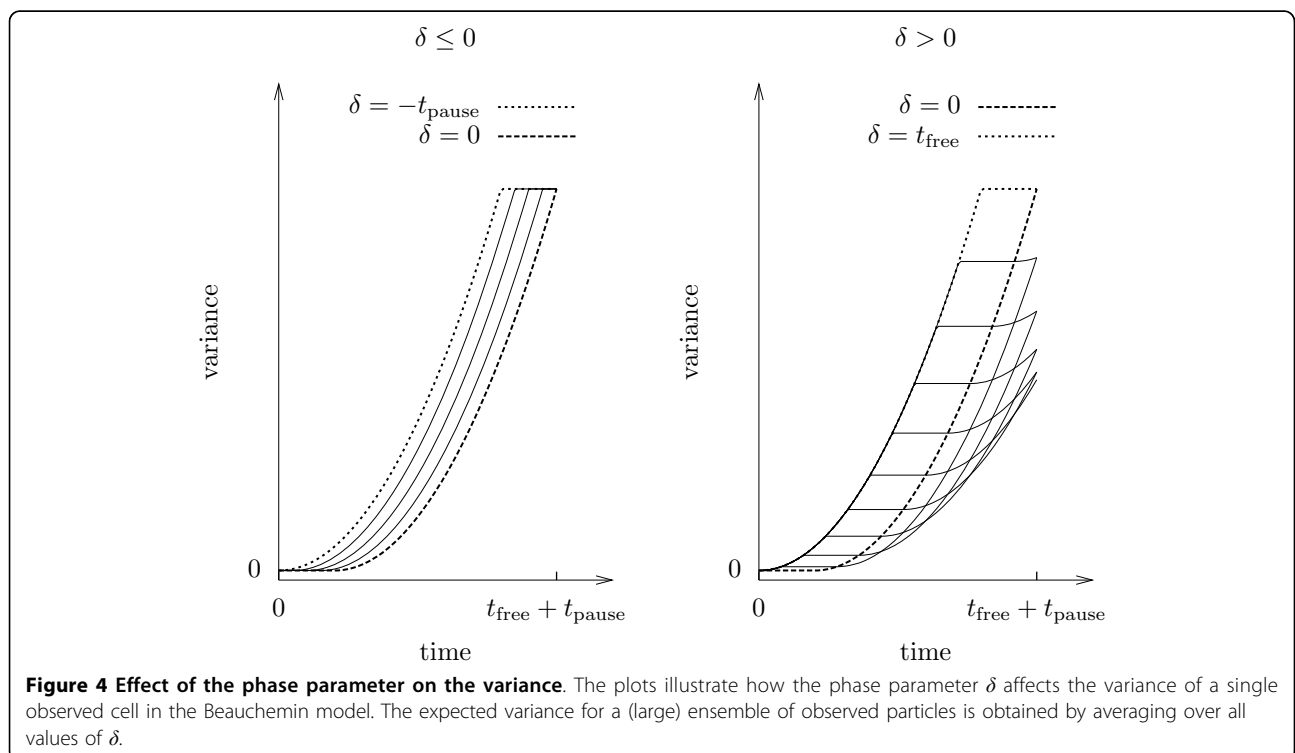
$$E[Var[\hat{x}(t)]] = \frac{1}{t_{free} + t_{pause}} \int_{\tau=0}^{t_{free}+t_{pause}} Var[x(t + \tau_0)] - Var[x(\tau_0)] d\tau_0$$

We start by integrating the two variance terms. Some algebra yields

$$\frac{1}{t_{free} + t_{pause}} \int_{\tau_0=0}^{t_{free}+t_{pause}} Var[x(t + \tau_0)] + Var[\tau_0] d\tau_0$$

$$= \frac{(v_{free} t_{free})^2}{3(t_{free} + t_{pause})} \left(\left(t + \frac{t_{free}}{3} \right) + \frac{t_{free}}{3} \right)$$

$$= 2M \left(\left(t + \frac{t_{free}}{3} \right) + \frac{t_{free}}{3} \right)$$

Figure 4 Effect of the phase parameter on the variance. The plots illustrate how the phase parameter δ affects the variance of a single observed cell in the Beauchemin model. The expected variance for a (large) ensemble of observed particles is obtained by averaging over all values of δ.

where M is the motility coefficient as defined in equation 2. For the covariance term it can be shown by some case distinctions that

$$\frac{1}{t_{\text{free}} + t_{\text{pause}}} \int_{\tau_0 = 0}^{t_{\text{free}} + t_{\text{pause}}} \text{Cov}[x(t + \tau_0), \tau_0] d\tau_0$$

$$= M \min(t, t_{\text{free}}) + M \underbrace{\left(\frac{\min(t, t_{\text{free}})^2}{t_{\text{free}}} + \frac{\min(t, t_{\text{free}})3}{3t_{\text{free}}^2} + \frac{2t_{\text{free}}}{3} \right)}_{=0 \text{ for } t \geq t_{\text{free}}}$$

Combining the variance and covariance terms as per expression 6, we obtain the claimed identity. □

As illustrated in Figure 5, the empirical mean square displacement from time t_{free} onwards is thus perfectly described by the linear equation

$$\text{E}[D^2(t)] = \alpha t + \beta \tag{7}$$

with slope

$$\alpha = 2M = \frac{(v_{\text{free}})^2 (t_{\text{free}})^2}{3(t_{\text{free}} + t_{\text{pause}})} \tag{8}$$

and intercept

$$\beta = -\frac{2t_{\text{free}}}{3} M = -\frac{(v_{\text{free}})^2 (t_{\text{free}})^3}{9(t_{\text{free}} + t_{\text{pause}})}. \tag{9}$$

This is an important difference to other models of mean square displacement such as Fürth's equation,

which is frequently applied to cell migration data [16,17]. This equation reads

$$\text{E}[D^2(t)] = 2M(t - P(1 - e^{-t/P})) \tag{10}$$

where P is a parameter called the *persistence length*. Hence, in the Beauchemin Model, the persistence of migration carries on only for the predefined time $t_{\text{pause}} + t_{\text{free}}$, after which it is immediately and completely lost. On the other hand, persistence in Fürth's equation gradually decays over time following an exponential term. This is why we previously used Fürth's equation to estimate the cell motility coefficient from short-term migration data [7]. However, whether this really leads to more accurate results remains to be determined.

A second notable consequence of the above result is that although the Beauchemin model has *three* microscopic parameters, its step-wise variance has only *two* degrees of freedom. Consequently, while the value of t_{pause} would become apparent when analyzing the track of a single simulated cell at a very high resolution, it does no longer matter when a large population of cells that migrate asynchronously is analyzed, as shown by Proposition 12: Given a fixed t_{free}, there are infinitely many combinations of t_{pause} and v_{free} that yield an identical motility coefficient and thereby an identical expected square displacement of the population. For example, the empirical mean square displacement of the best-fitting triplet (0.5, 2.0, 18.8) is identical (up to some
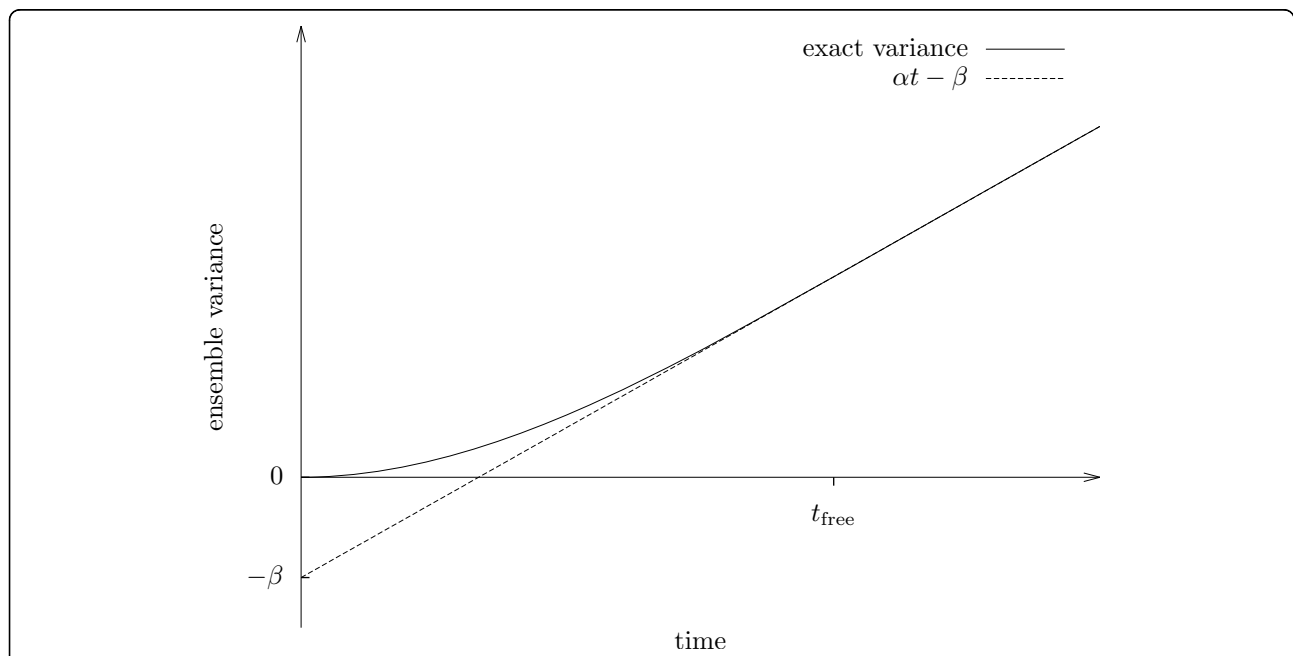

Figure 5 Forming the population average. Expected mean square displacement of observed particle ensembles as of Proposition 12. The pulses of the individual variances (see Figure 4) average out, and the exact expression converges quickly to a linear form with slope $\alpha = 2M$ and intercept $\beta = 2Mt_{\text{free}}/3$. One degree of freedom from the individual variance is lost due to the averaging process.

rounding) to that of the triplet (9.32, 2.0, 40). Hence, we cannot expect to get reliable estimates for all three parameters $t_{\text{free}}, t_{\text{pause}}$ and v_{free} from fitting the model to mean square displacement data. This is consistent with the observation by Beauchemin et al. that the parameter t_{pause} can always be set to 0 without much influence on the quality of the best fit.

However, Beauchemin et al. fitted their simulations to the *mean displacement* data, and the above result does not imply that the mean displacement too has just two degrees of freedom. Perhaps surprisingly, simulations show that this is in fact not the case (Figure 6): While mean square displacement plots generated with using the above parameter triplets and our analytical solutions are indeed consistent with each other, the mean displacement plots are significantly different. Hence, the mean displacement plot appears to preserve more information on the underlying migration process than does the mean square displacement plot, at least when applied to cell tracks simulated using the Beauchemin model.

Confinement ratio

The confinement ratio is a measure of track straightness. In the context of lymphocyte migration analysis, it is also called "chemotactic index" or "meandering index" [2,3]. The confinement ratio is usually defined as the quotient of a path's displacement from the origin versus the complete length of the path. If the cell changes directions very often, the confinement ratio will tend to zero over time, while for directed

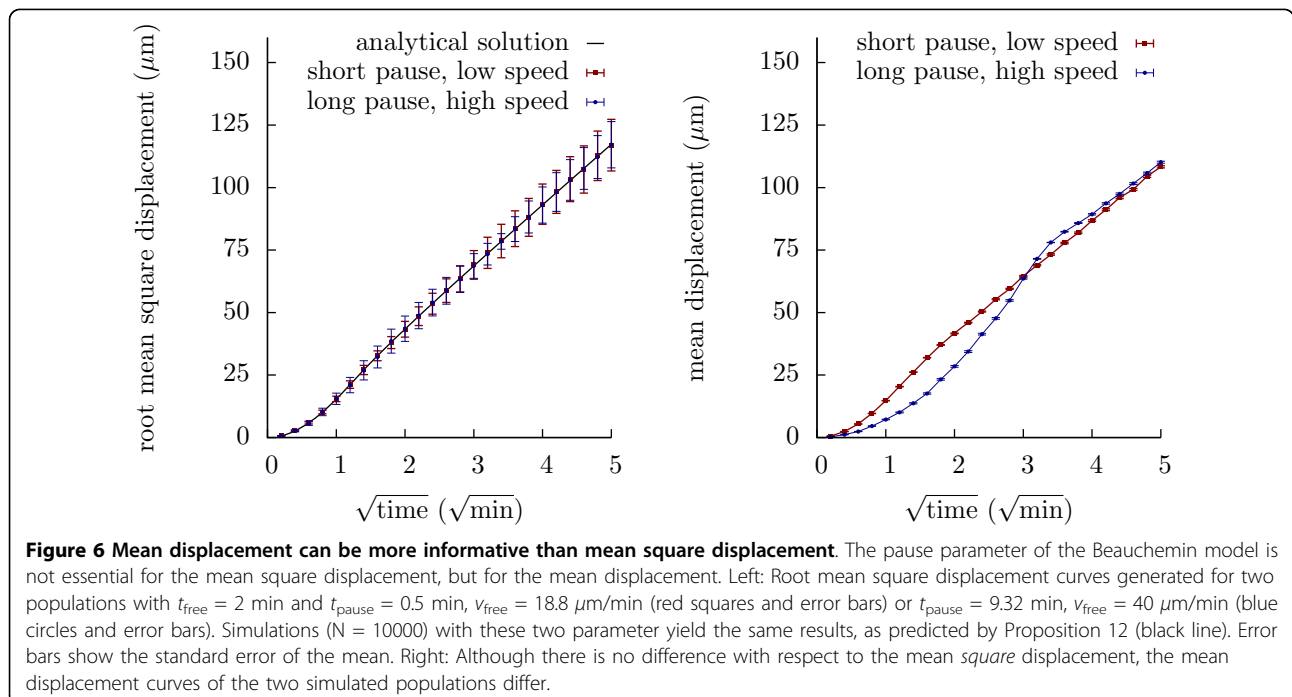migration without a random component, the confinement ratio is equal to 1.

Instead of the confinement ratio, we consider here the *squared* confinement ratio, which can be easily determined for the Beauchemin model from the previous result.

Definition 13 (Squared confinement ratio). *Consider a particle moving according to the Beauchemin model and let $D^2(t)$ denote the squared distance between the particle position at time t and its starting point. Then we define the squared confinement ratio $C^2(t)$ as follows:*

$$C^2(t) = \frac{D^2(t)}{t^2 v_{\text{free}}^2}$$

By assuming $t_{\text{pause}} = 0$ and applying the results from the previous section, we can determine the expected confinement ratio of the Beauchemin model as follows:

$$\mathrm{E}\left[C^2(t)\right] = \frac{1}{t^2 v_{\text{free}}^2}\left[2Mt - 2Mt_{\text{free}} \times \begin{cases} \frac{1}{3} & t \geq t_{\text{free}} \\ \frac{1}{3}\left(\frac{t}{t_{\text{free}}}\right)^3 - \left(\frac{t}{t_{\text{free}}}\right)^2 + \left(\frac{t}{t_{\text{free}}}\right) & t < t_{\text{free}} \end{cases}\right]$$

$$= \frac{1}{t^2}\left[\frac{t_{\text{free}}}{3}t - \frac{t_{\text{free}}^2}{3} \times \begin{cases} \frac{1}{3} & t \geq t_{\text{free}} \\ \frac{1}{3}\left(\frac{t}{t_{\text{free}}}\right)^3 - \left(\frac{t}{t_{\text{free}}}\right)^2 + \left(\frac{t}{t_{\text{free}}}\right) & t < t_{\text{free}} \end{cases}\right]$$

$$= \frac{t_{\text{free}}}{3t} - \frac{t_{\text{free}}^2}{3t^2} \times \begin{cases} \frac{1}{3} & t \geq t_{\text{free}} \\ \frac{1}{3}\left(\frac{t}{t_{\text{free}}}\right)^3 - \left(\frac{t}{t_{\text{free}}}\right)^2 + \left(\frac{t}{t_{\text{free}}}\right) & t < t_{\text{free}} \end{cases}$$

$$= \frac{t_{\text{free}}}{3t} - \frac{1}{9} \times \begin{cases} 1 & t \geq t_{\text{free}} \\ \frac{t}{t_{\text{free}}} - 3 + \frac{3t_{\text{free}}}{t} & t < t_{\text{free}} \end{cases}$$

$$= \frac{t_{\text{free}}}{3t} - \frac{1}{9} \times \begin{cases} 1 & t \geq t_{\text{free}} \\ \frac{3t_{\text{free}}^2 + t^2}{t_{\text{free}}t} - 3 & t < t_{\text{free}} \end{cases}$$

$$= \cdots$$

$$= \frac{1}{3\max(t, t_{\text{free}})}\left(t_{\text{free}} - \frac{t \times t_{\text{free}}}{3\max(t, t_{\text{free}})^2}\right)$$

Figure 6 Mean displacement can be more informative than mean square displacement. The pause parameter of the Beauchemin model is not essential for the mean square displacement, but for the mean displacement. Left: Root mean square displacement curves generated for two populations with $t_{\text{free}} = 2$ min and $t_{\text{pause}} = 0.5$ min, $v_{\text{free}} = 18.8$ μm/min (red squares and error bars) or $t_{\text{pause}} = 9.32$ min, $v_{\text{free}} = 40$ μm/min (blue circles and error bars). Simulations (N = 10000) with these two parameter yield the same results, as predicted by Proposition 12 (black line). Error bars show the standard error of the mean. Right: Although there is no difference with respect to the mean *square* displacement, the mean displacement curves of the two simulated populations differ.

Note that $E[C^2(0)] = \frac{1}{3}$, because the path length tv_{free} is measured in three dimensions while the expected observed square distance is projected to one dimension.

Beltman et al. [3] have argued that the confinement ratio is not well suited for comparisons between different experiments because it tends to zero over time, and is therefore sensitive to the duration of the experiment from which it was estimated. They suggested to instead normalize the confinement ratio by multiplying it with the square root of time, such that it converges to a constant value. Applying this normalization to the squared confinement ratio of the Beauchemin model, we obtain

$$t \times E[C^2(t)] = \frac{t}{3\max(t, t_{\text{free}})}\left(t_{\text{free}} - \frac{t \times t_{\text{free}}}{3\max(t, t_{\text{free}})^2}\right), \quad (11)$$

which for $t \geq t_{\text{free}}$ simplifies to

$$t \times E[C^2(t)] = \frac{t_{\text{free}}(3t - 1)}{9t}. \quad (12)$$

Thus, the normalized squared confinement ratio converges to $t_{\text{free}}/3$. When there is evidence that the observed migration process really behaves roughly like the Beauchemin model, this could be a good rule of thumb for quick estimations of the persistence time from experimental data.

Simulating biased migration

For our recent study of the detection limits of two-photon imaging [7], we extended the Beauchemin model for simulating three different types of biased migration, for which we used the term *taxis modes* [18] to indicate that such biased migration in cells usually occurs in response to an external stimulus. These three taxis modes were *orthotaxis*, where the movement speed is adjusted, *klinotaxis*, where the duration of the "free runs" is changed, and *topotaxis*, where the turning angle distribution of the simulated cell is changed. Thus, these modifications can be viewed as semi-mechanistic ones that attempt to describe ways in which a cell could re-adjust its migration behaviour from pure random migration to biased random migration in response to an external stimulus. Before we review these modifications, we start by presenting a fourth method to simulate biased migration which is not mechanistically motivated, but is much easier to analyze.

In all cases, taxis is spatially and temporarily uniform and controlled by two parameters: The vector $\vec{b} \in \mathbb{R}^3$, assumed to have unitary length, sets the direction of the taxis, and the parameter $p \in [0, 1]$ determines the magnitude of taxis, with $p = 0$ equivalent to an unbiased random walk and $p = 1$ to a maximally biased random walk; the precise meaning of the parameter p will be defined for each case below.

Simple phenomenological model

A simple ad-hoc method to extend the Beauchemin model for simulating biased migration is the following: During the pause phase, the cell is no longer kept stationary, but moves into the bias direction $\vec{b}$ with speed $p \cdot v_{\text{free}}$. This model is very simple to analyze, because the modification is deterministic and so changes only the mean, but not the variance of each step. Hence, the motility coefficient remains the same.

Proposition 14 (Taxis speed for the simple model). *The length of the convection coefficient, $||C||$, of the simple biased migration model is given by*

$$||C|| = \frac{p \cdot v_{\text{free}} \cdot t_{pause}^2}{t_{\text{free}} + t_{pause}}.$$

Proof. During each pause phase, the cell moves a distance $p \cdot v_{\text{free}} \cdot t_{pause}$ into the bias direction. $||C||$ is then simply the product of this distance and the total fraction of time that the cell spends in the pause phase. □

There might be some use for the simple biased migration model in situations where it is necessary to study the effect of bias independently of the motility coefficient. However, it is difficult to imagine how such a situation could arise biologically. Below we therefore discuss model modifications that have a more realistic motivation. For each of these modifications, it is possible to analytically derive the convection coefficient, i.e., the expected speed of the bias.

Throughout the rest of this section, we let the vector $\vec{d}$ denote the orientation vector of the Beauchemin model (i.e., a vector randomly sampled from the unit sphere), and $\vec{b}$ the bias direction. Moreover, we let (ξ', v', ζ') denote the vector of random variables representing a step of the modified model, while (ξ, v, ζ) denotes a step in the original model.

Orthotaxis model

In the orthotaxis model, the migration speed is no longer constant but depends on the chosen target direction. The more consistent this direction is with the target direction, the faster the simulated cell will move. Such bias could occur for instance in response to some external dragging force, and might be illustrated by imagining a random stroll through a city on a very windy day.

Formally, we define the adjusted speed as follows:

$$v'_{\text{free}} = v_{\text{free}}\left(1 + p\langle\vec{b}, \vec{d}\rangle\right)$$

Here $\langle\cdot,\cdot\rangle$ denotes the scalar product and p the bias strength parameter. Note that the movement perpendicular to $\vec{b}$ is not affected at all, while movement against

the direction $\vec{b}$ is slower and in direction $\vec{b}$ it is faster. More precisely, the relation between p and the bias speed v_{taxis} is given by the following proposition.

Proposition 15 (Speed of orthotaxis [7]). *The convection coefficient $\|C\|$ of the orthotaxis model is given by*

$$\|C\| = v_{\text{free}} \cdot \frac{p \, t_{\text{free}}}{3(t_{\text{free}} + t_{\text{pause}})}.$$

Proof. We can assume without loss of generality that the bias is along the first dimension, i.e. $\vec{d} = (1, 0, 0)$, and thus $\langle \vec{b}, \vec{d} \rangle = \xi$. We are interested in the expectation $E[\xi']$; it is easy to see that the expectations $E[v']$ and $E[\zeta']$ are still 0. Then we get

$$\xi' = v'_{\text{free}}\xi = v_{\text{free}}(1 + p\xi)\xi$$
$$= v_{\text{free}}(\xi + p\xi^2).$$

Because ξ is uniformly distributed on $[-1, 1]$ (see proof of Proposition 10), the expected location $E[\xi']$ at the end of the step is

$$E[\xi'] = E[v_{\text{free}} \cdot t_{\text{free}}(\xi + p\xi^2)]$$
$$= v_{\text{free}} \cdot t_{\text{free}}(E[\xi] + pE[\xi^2])$$
$$= v_{\text{free}} \cdot t_{\text{free}} \cdot p/3.$$

Dividing this expected location $E[\xi']$ by the step duration $t_{\text{free}} + t_{\text{pause}}$ yields the claimed expression for $\|C\|$. □

Topotaxis model

In the case of topotaxis, the cell adjusts its turning behaviour but keeps the speed constant. This also affects both the variance and the mean of the cell step distribution. Topotaxis could be a likely bias mechanism for cells migrating on some sort of structural cell network in the tissue, like it was suggested to be the case for T cells in secondary lymphoid organs [19].

We implement topotaxis by skewing the distribution from which we pick the cell's orientation $\vec{d} = (\xi, v, \zeta)$. Recall that this is a uniform distribution on a sphere in the unbiased case. Such a distribution can also be generated by sampling ξ uniformly from $[-1, 1]$ and then picking v, ζ uniformly from a circle with radius $\sqrt{1 - \xi^2}$. Thus, the probability density function $f_\xi(x)$ is defined as follows:

$$f_\xi(x) = \frac{1}{2}, \quad x \in [-1, 1].$$

For the modified model version, we define a parameterized version of $f_\xi(x)$ as follows:

$$f_{\xi'}(x) = \frac{1 + px}{2}, \quad x \in [-1, 1].$$

Hence, for $p = 0$ we obtain the distribution $f_{\xi'}(x) = $ ⟋ of the unbiased case, while for $p = 0.5$ we get the slightly skewed distribution $f_{\xi'}(x) = $ ⟋ and for $p = 1$ we obtain the maximally skewed distribution $f_{\xi'}(x) = $ ⟋.

This modification of the Beauchemin model yields the same convection coefficient as in the previous case.

Proposition 16 (Speed of topotaxis [7]). *The speed $\|C\|$ of the directional motion induced by the topotaxis model is given by the equation*

$$\|C\| = v_{\text{free}} \cdot \frac{p \, t_{\text{free}}}{3(t_{\text{free}} + t_{\text{pause}})}.$$

Proof. This is simply shown by calculating $E[\xi']$ for the probability density function defined above. □

Klinotaxis model

Lastly, we can also induce bias by modifying the duration t_{free} of the free run depending on the picked orientation $\vec{d}$. A well-known biological example of such bias is the "run-and-tumble" motion by which *E. Coli* searches for food: Because the bacterium is too small to sense concentration gradients *in situ*, it is capable of comparing concentration measurements along a persistent path until it figures out whether it is going in the right direction. When it senses that it is moving in the right direction (run), it will try to keep going, while otherwise it will stop its motion soon (tumble) and orient itself in a new random direction.

We implement klinotaxis by modifying the free run duration t_{free} depending on the angle between $\vec{d}$ and $\vec{b}$ as follows:

$$t'_{\text{free}} = t_{\text{free}}(1 + p\langle \vec{b}, \vec{d} \rangle)$$

Again, this modification leads to exactly the same bias speed as found for the previous two models.

Proposition 17 (Speed of klinotaxis [7]). *The speed $\|C\|$ of the directional motion induced by the klinotactic modification of the Beauchemin model is given by the equation*

$$\|C\| = v_{\text{free}} \cdot \frac{p \, t_{\text{free}}}{3(t_{\text{free}} + t_{\text{pause}})}.$$

Proof. The proof follows along the lines of a similar model of *E. Coli* run-and-tumble motion discussed by de Gennes [20]. We start again by analyzing a single free run of a cell projected onto the direction of bias, which is assumed without loss of generality to be along the first dimension ξ. Again we need to determine $E[\xi']$. We start by writing ξ' as

$$\xi' = v_{\text{free}} \cdot \xi \cdot t_{\text{free}} \cdot (1 + p\xi).$$

Now let us assume for a moment that $t_{pause} = 0$. It is easy to verify that $E[t'_{free}] = t_{free}$. Thus, the average speed towards the bias direction over many successive steps is $E[\xi'/t_{free}]$. A simular calculation as in the proof of Proposition 15 yields

$$E[\xi'/t_{free}] = v_{free} \cdot p/3.$$

Now we arrive at the claimed expression by multiplying the above with the overall fraction of time that the particle spends in the free run phase, which is

$$t_{free}/(t_{free} + t_{pause}).$$

□

Note that the klinotaxis model no longer fits within our mathematical framework because the duration of a step in the model is no longer constant. Thus, the central limit theorem we applied in that section no longer applies, and a different central theorem would be needed to show formally that the model converges to a Brownian motion.

Moreover, we point out that for all modifications discussed in this section except for the simple phenomenological one, the random motility component changes. Therefore, these modifications cause as a side-effect a modified motility coefficient. In general, the random motility component are even affected in a non-isotropic fashion, such that the motility coefficient would need to be described using a 3 × 3 matrix instead of a single scalar value. In Figure 7, we show some simulation results that illustrate this point.
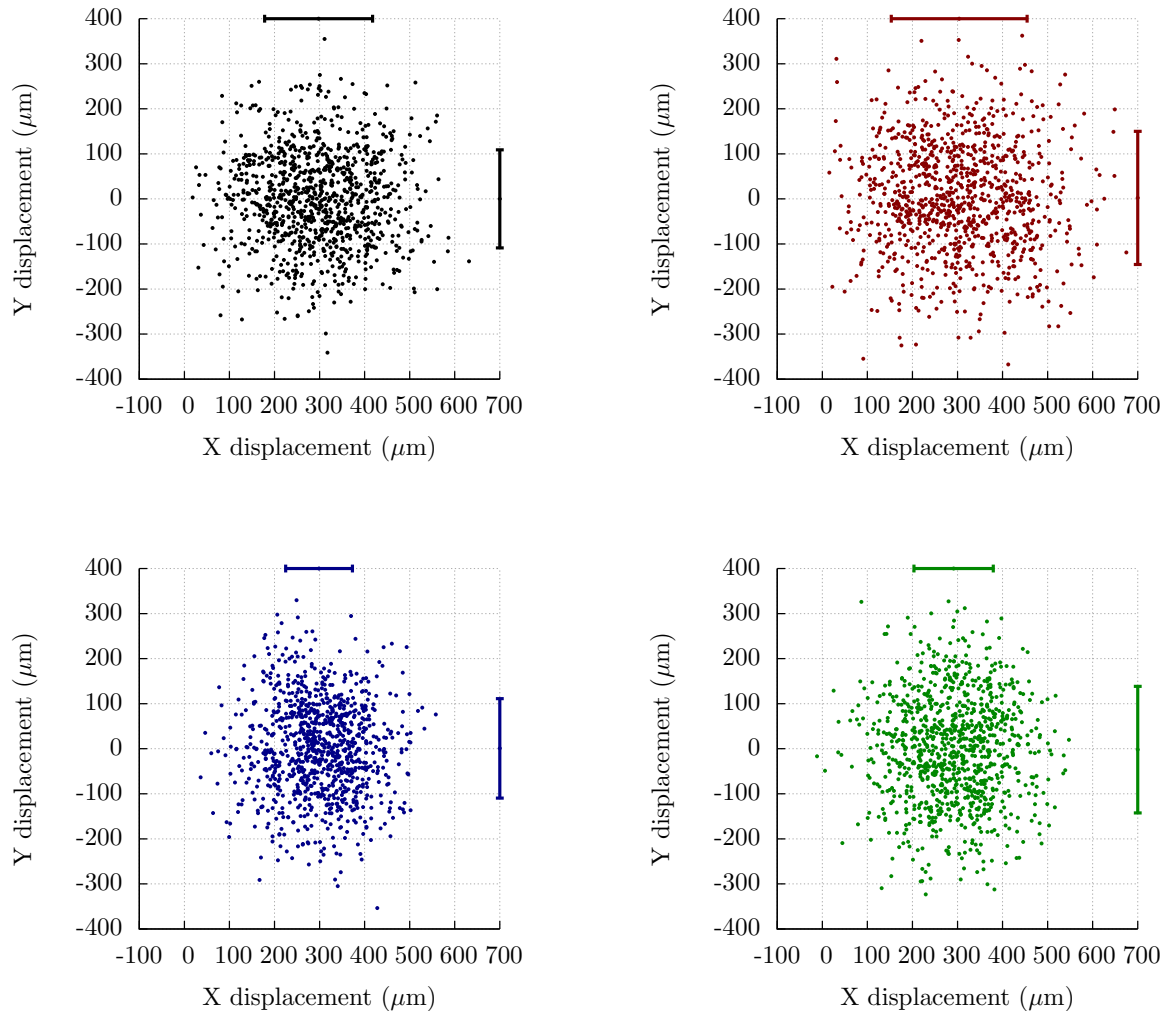

Figure 7 Taxis skews the random motility component. Displacement of *in silico* cells when applying the four different methods of simulating biased migration along the X axis using the methods discussed in this paper, i.e., the simple phenomenological model (black), orthotaxis (red), topotaxis (blue), and klinotaxis (green). For each panel, $N = 1000$ cells were simulated using the parameters $t_{pause} = 0.5$ min, $t_{free} = 2.0$ min, and $v_{free} = 18.8$ μm/min. Cell displacements are shown after one hour of simulation time, and the taxis parameters p are in each case set such that the expected displacement is 300 μm per hour. The length of the scale bars is proportional to the dimension-wise variances of the cell displacements, demonstrating that these variances change when biased migration is introduced.

Conclusions

In this article, we have compiled both old and new analytical results on the Beauchemin model of lymphocyte migration, which was originally a model of purely random motility, but is now able to accommodate partially biased migration as well. An important practical consequence of these results is that validation of the model against experimentally determined motility coefficient, mean square displacement data, and confinement ratio curves does no longer need to be performed by simulation, because these quantities can be calculated directly from the model parameters. This also allows for the use of standard numerical fitting procedures, such as implemented in popular computer algebra packages, to estimate the model parameters from such data by least-squares-fitting. An important observation was that when simulating a large population of unsynchronized cells, the model is equivalent to the ideal chain model of polymer physics, which is however only true with respect to the quantities named above and not for others like the mean displacement. Interestingly, this shows that we can sometimes infer information from mean displacement data that we cannot infer from mean *square* displacement data. However, an analytical derivation of the mean displacement of the model has not yet been achieved, and can be expected to be complicated judging from the experiences with similar models [21]. For now, we leave this as an important open problem.

Subtle biased migration of lymphocytes can point to important biological functions or phenomena [22,23], and the use of mathematical models has helped to understand the quantitative impact of such effects [23,7]. The modifications discussed above facilitate simulations of such subtle biases with pre-defined speed, again removing the need for fitting the bias strength parameter to data by using simulations. However, all but one of the discussed modifications also affect the random migration component, which is likely also true in biological scenarios where randomly migrating cells respond to an external stimulus. Therefore, another important task that remains to be solved is to analytically derive the motility coefficient matrices for the topotaxis, klinotaxis, and orthotaxis models.

Authors' contributions

JT conceived the study, participated in the mathematical analysis, and wrote the paper; MS participated in the mathematical analysis and helped write the paper; RJdB participated in the study conception and helped write the paper.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Theoretical Biology & Bioinformatics, Utrecht University, 3508 TC Utrecht, The Netherlands. [2]Institute for Theoretical Computer Science, University of Lübeck, 23562 Lübeck, Germany. [3]IBM Research, Mulhuddart, Dublin 15, Ireland.

References

1. Miller MJ, Wei SH, Parker I, Cahalan MD: Two-photon imaging of lymphocyte motility and antigen response in intact lymph node. *Science* 2002, **296**:1869-1873.
2. Cahalan MD, Parker I: Choreography of cell motility and interaction dynamics imaged by two-photon microscopy in lymphoid organs. *Annual Reviews in Immunology* 2008, **26**:585-626.
3. Beltman JB, Marée AFM, de Boer RJ: Analysing immune cell migration. *Nature Reviews Immunology* 2009, **9**:789-798.
4. Beltman JB, Marée AFM, Lynch JN, Miller MJ, de Boer RJ: Lymph node topology dictates T cell migration behavior. *Journal of Experimental Medicine* 2007, **204**:771-780.
5. Beauchemin C, Dixit NM, Perelson AS: Characterizing T cell movement within lymph nodes in the absence of antigen. *Journal of Immunology* 2007, **178(9)**:5505-5512.
6. Grigorova IL, Panteleev M, Cyster JG: Lymph node cortical sinus organization and relationship to lymphocyte egress dynamics and antigen exposure. *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107**:20447-20452.
7. Textor J, Peixoto A, Henrickson SE, Sinn M, von Andrian UH, Westermann J: Defining the Quantitative Limits of Intravital Two-Photon Lymphocyte Tracking. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108(30)**:12401-12406.
8. Flory PJ: *Statistical Mechanics of Chain Molecules* Interscience Publishers; 1969.
9. Berg HC: *Random Walks in Biology* Princeton University Press; 1993.
10. Kallenberg O: *Foundations of modern probability* Springer; 2002.
11. Murray JD: *Mathematical Biology: I. An Introduction.* 3 edition. Springer; 2002.
12. Billingsley P: *Convergence of Probability Measures.* 2 edition. Wiley, John & Sons; 1999.
13. Evans LD: *Partial Differential Equations* American Mathematical Society; 1998.
14. Donovan GM, Lythe G: T-cell movement on the reticular network. *Journal of Theoretical Biology* 2012, **295**:59-67.
15. Marsaglia G: Choosing a point from the surface of a sphere. *Annals of Mathematical Statistics* 1972, **43(2)**:645-646.
16. Selmeczi D, Mosler S, Hagedorn PH, Larsen NB, Flyvbjerg H: Cell Motility as Persistent Random Motion: Theories from Experiments. *Biophysical Journal* 2005, **89**:912-931.
17. Grabher C, Cliffe A, Miura K, Hayflick J, Pepperkok R, Rørth P, Wittbrodt J: Birth and life of tissue macrophages and their migration in embryogenesis and inflammation in medaka. *Journal of Leukocyte Biology* 2007, **81**:263-271.
18. Ionides EL, Fang KS, Isseroff RR, Oster G: Stochastic models for cell motion and taxis. *Journal of Mathematical Biology* 2003, **48**:23-37.
19. Bajénoff M, Egen JG, Koo LY, Laugier JP, Brau F, Glaichenhaus N, Germain RN: Stromal cell networks regulate lymphocyte entry, migration, and territoriality in lymph nodes. *Immunity* 2006, **25**:989-1001.
20. de Gennes PG: Chemotaxis: the role of internal delays. *European Biophysics Journal* 2004, **33**:691-693.
21. Codling EA, Plank MJ, Benhamou S: Random walk models in biology. *Journal of the Royal Society Interface* 2008, **5(25)**:813-834.
22. Castellino F, Huang AY, Altan-Bonnet G, Stoll S, Scheinecker C, Germain RN: Chemokines enhance immunity by guiding naive CD8+ T cells to sites of CD4+ T cell-dendritic cell interaction. *Nature* 2006, **440**:890-895.

23. Beltman JB, Allen CD, Cyster JG, de Boer RJ: **B cells within germinal centers migrate preferentially from dark to light zone.** *Proceedings of the National Academy of Sciences of the United States of America* 2011.

24. Mempel TR, Henrickson SE, von Andrian UH: **T-cell priming by dendritic cells in lymph nodes occurs in three distinct phases.** *Nature* 2004, **427**:154-159.

25. Miller MJ, Hejazi AS, Wei SH, Cahalan MD, Parker I: **T cell repertoire scanning is promoted by dynamic dendritic cell behavior and random T cell motility in the lymph node.** *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101**:998-1003.

26. Miller MJ, Wei SH, Cahalan MD, Parker I: **Autonomous T cell trafficking examined** *in vivo* **with intravital two-photon microscopy.** *Proceedings of the National Academy of Sciences of the United States of America* 2003, **100**(5):2604-2609.