# From random walks to fully anisotropic diffusion models for cell and animal movement

Kevin J. Painter and Thomas Hillen

**Abstract** This chapter provides an introduction on how anisotropic diffusion models can be derived from position-jump and velocity-jump random walks. We show how the availability of measurement data can guide the choice of the appropriate model. We further present two new applications, respectively to cell movement on micro-fabricated surfaces and magnetic compass orientation by sea turtle hatchlings.

## 1 Introduction

Getting from point A to point B is a daily challenge, although for the most part our movement patterns are routine – staggering from bedroom to bathroom, from home to work, from office to coffee pot – and we switch into autopilot, following the course hard-wired into our conscious. Sometimes we may find ourselves in an unusual place attempting to reach an unfamiliar goal, yet even then navigation is straightforward when armed with a smartphone and network connection.

Cells and animals do not have the technological aids at our disposal yet frequently need to migrate through their environment, sometimes independently, sometimes collectively: the solo navigations of recently fledged albatrosses across thousands of kilometres of southern oceans, or the collective movements of cells as they move into developing tissues and organs offer particularly astonishing examples. Given the myriad of potential factors – chemicals, electric, magnetic and gravitational fields, topography and physical structure of the environment, etc – a key question, whether posed by ecologists, cell biologists, microbiologists or oncologists, is exactly what cues signal to the cells or organisms along their paths.

Mathematical and computational modelling offer the means to address such questions, via encapsulating a biological process into its essentials. Yet choosing an approach and setting up a model to begin with is far from a trivial task. Inevitably this will come down to the knowledge and data we have and the nature of

K. J. Painter
Department of Mathematics and Maxwell Institute for Mathematical Sciences, Heriot-Watt University, UK & Dipartimento di Scienze Matematiche, Politecnico di Torino, Italy, e-mail: K.Painter@hw.ac.uk

T. Hillen
University of Alberta, Edmonton, Canada e-mail: thillen@ualberta.ca

the problem we are trying to address. One major determinant in the modelling choice will be the *biological scale* of the problem. Consider a population-scale problem such as predicting the spatial spread of a cancer to aid diagnosis and treatment. While we may have some understanding of the underlying biological processes at a cellular level (e.g. enhanced proliferation and invasion of cells into healthy tissue), the primary scale of interest is typically a macroscopic one at the time of treatment: the scale of the cancer (centimetres) is significantly greater than the microscopic cells from which it is formed. In such instances, an efficient and oft-used solution is to blur the population into a convenient density distribution and propose a suitable evolution equation (such as a partial differential equation) for its change over space and time [48, 37, 36].

Macroscopic approaches such as these have formed a bedrock for mathematical modelling over many years, providing insight into a wide variety of fundamental processes. When the only data we have is similarly macroscopic, such as an MRI (magnetic resonance imaging) scan indicating the spatial extent of a cancer's growth, a macroscopic model makes sense: fitting the model to approximated densities determined from the scan offers a method of validation and parameter estimation [56]. But what if the available data is at the level of the individual? Can we relate a model posed at a macroscopic level to an individual's movement? These questions are clearly crucial when we consider technological advances in our capacity to track molecules, cells or organisms: individual molecules can be tagged and followed via single particle tracking (SPT) as they skate across the cell membrane [52]; labelled cells can be followed via sophisticated imaging while migrating through a complicated tissue environment [59]; attaching a global positioning system (GPS) to an animal can allow it to be followed even if it travels across oceans and continents [7]. Clearly, the data provided by such methods can shed significant light on the fundamental mechanisms of movement. For modellers, a significant challenge is raised: how can we best exploit all forms of available data to obtain better models, both at the level of individuals and populations?

To motivate the rest of this chapter, we consider two very different applications respectively in cell movement and turtle hatchling navigation. Both applications have a similar fundamental question (what are the guidance cues that determine navigation?), but offer distinct examples for the type of data that may be at hand for model parametrisation/formulation. In the case of cell movement we have a tabulated summary of population-averaged behaviour. For turtles we have individual-level data, an orientation for each tested hatchling in a sample. The analytical models we proceed to describe can be fitted to each of the datasets, in each case shedding light on the problem.

## *1.1 Dataset A: Cell Movement on Microfabricated Substrates*

The development, maintenance and repair of our bodies requires that various cells migrate through complex tissue environments; in tumour invasion, these same mechanisms can facilitate the rapid dispersal and spread of malignant cells into neighbouring healthy tissue [19]. Various extracellular factors contribute to cell guidance, ranging from extracellular molecules (e.g. chemoattractants and repellents), direct signals from other cells (e.g. contact inhibition of locomotion) and the oriented movement of cells along aligned structures [20, 42]. This latter form of oriented movement is generally termed contact guidance [14] and, while principally described in the context of movement along the long bundles of collagen fibres characteristic of connective tissue, can also occur during the movement of cells along axonal tracts of the central nervous system or crawling along blood capillaries [17]. Contact guidance has been identified in various cell

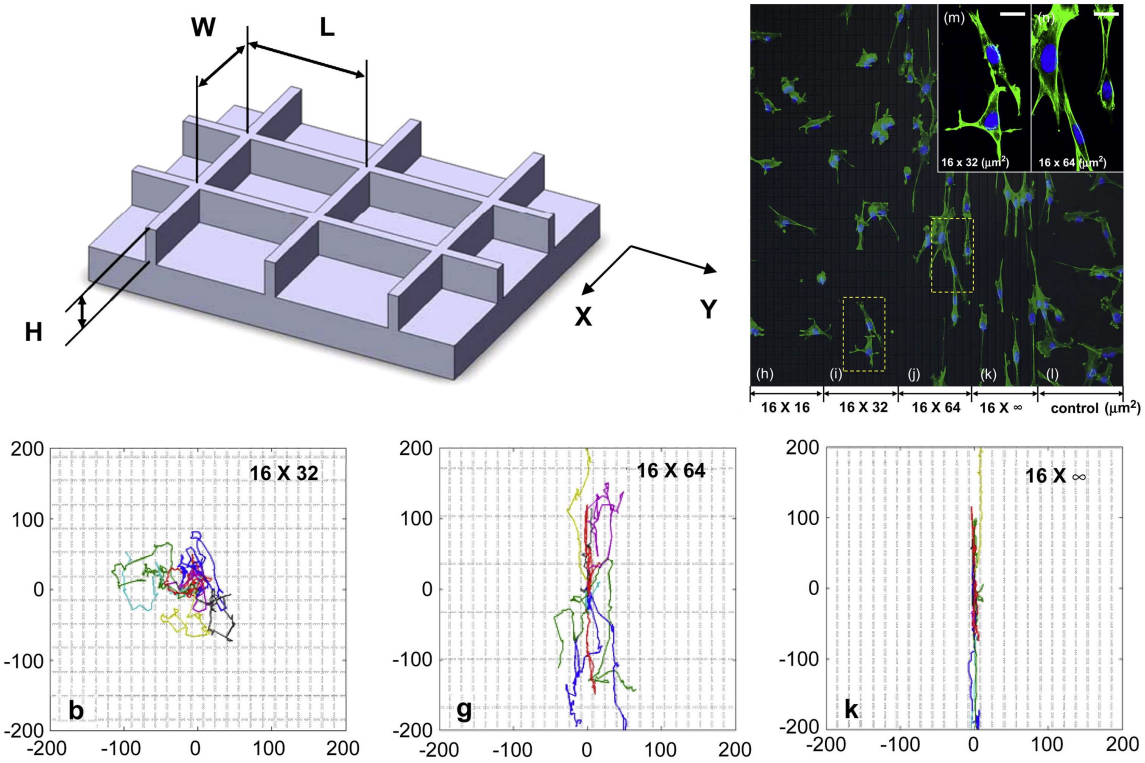


**Fig. 1** Top Left: schematic of the micro-ridge substrate. Top Right: typical observation of cell movement on an anisotropic substrate, where the micro-ridges are in different aspect ratios. Bottom: cell tracks observed for different environmental anisotropies. Horizontal and vertical axes represent microns. Figures reprinted from Biomaterials, volume 31, Jeon, H., Hidai, H., Hwang, D.J., Healy, K.E. and Grigoropoulos, C.P., "The effect of micronscale anisotropic cross patterns on fibroblast migration", pp. 4286–4295 (2010), with permission from Elsevier.

| Case ($\mu$m$\times\mu$m) | Ridge height ($\mu$m) | $x$-velocity $v_x\pm$error ($\mu$m/min) | $y$-velocity $v_y\pm$ error ($\mu$m/min) | Speed $\pm$ error ($\mu$m/min) |
|---|---|---|---|---|
| 12 x 24 | 3 | 0.38±0.015 | 0.58±0.025 | 0.78±0.027 |
| 12 x 48 | 3 | 0.28±0.014 | 0.9±0.045 | 1.01±0.045 |
| 12 x ∞ | 3 | 0.08±0.005 | 0.56±0.029 | 0.59±0.029 |
| 16 x 32 | 3 | 0.48±0.021 | 0.65±0.026 | 0.9±0.03 |
| 16 x 64 | 3 | 0.31±0.015 | 0.87±0.038 | 1.0±0.039 |
| 16 x ∞ | 3 | 0.12±0.007 | 0.8±0.036 | 0.84±0.036 |
| 24 x 48 | 3 | 0.26±0.015 | 0.42±0.024 | 0.55±0.027 |
| 24 x 96 | 3 | 0.2±0.012 | 0.49±0.02 | 0.58±0.022 |
| 24 x ∞ | 3 | 0.12±0.007 | 0.48±0.027 | 0.52±0.028 |
| 12 x 24 | 10 | 0.33±0.016 | 0.46±0.024 | 0.65±0.026 |
| 12 x 48 | 10 | 0.18±0.013 | 0.76±0.044 | 0.83±0.046 |
| 12 x ∞ | 10 | 0.04±0.003 | 0.60±0.032 | 0.61±0.032 |
| control | 0 | 0.38±0.019 | 0.41±0.033 | 0.63±0.025 |

**Table 1** Reproduction of the movement data from Jeon et al. [25] for fibroblast cells migrating on a micro-ridged substratum.

populations, including fibroblasts [13], immune cells [59] and various cancerous populations [49, 16].

The capacity of environmental anisotropy to influence cell orientation/movement can be studied by tracing cell paths when plated on micro-fabricated structures. To illustrate the data available from such experiments we analyse those in Jeon et al. [25], where a two-dimensional substratum is formed with a rectangular array of orthogonal micro-ridges, see Figure 1 (left). Inter-ridge lengths in the $x-$ and $y-$directions are respectively denoted $W$ and $L$, with the former set at $12, 24$ or $48$ $\mu$m and the latter set to generate $W : L$ ratios of $1 : 2$, $1 : 4$ or $1 : \infty$ (the last case corresponding to an absence of ridges in the $x-$direction). Ridge heights were set at $3$ $\mu$m, with further tests conducted at $10\mu$m and a control case without any ridges. NIH373 fibroblast cells were plated on these substrates: a population characterised by its mesenchymal movement with cells extending long protrusions to probe the environment. Cells clearly align to the micro-ridges, generating anisotropic movement (see Figure 1, top right and bottom row) under anisotropic arrangements. Data from individual tracking was summarised at a macroscopic level (averaged over the population) in terms of mean speeds and directional bias, reproduced in Table 1. In Section 4.1 we will use this data to parametrise an anisotropic diffusion model that describes cell spread for different anisotropies in the substratum.

## *1.2 Dataset B: Magnetic Navigation in Loggerhead Hatchlings*

Maritime navigation is undeniably hazardous. The frequent lack of visible landmass, turbulent currents and dramatic meteorological conditions resulted in frequent positional misreckoning (and shipwrecking) during the early ages of maritime traffic, stimulating governments of the time to propose prizes for a method of accurately establishing longitudinal coordinates. John Harrison's marine chronometer marked a pivotal moment in the transition towards (relatively) safe navigation [53]. Marine animals, of course, do not rely on such aids but many species routinely undertake long marine journeys [29], with one of the most phenomenal belonging to the loggerhead turtle (*Caretta caretta*). North Atlantic loggerhead hatchlings dash to the ocean from eggs laid at various nesting beaches and undergo a period of "frantic" swimming that transports them from the dangerous coastal waters to ocean circulatory currents such as the Gulf Stream. They subsequently embark on a years to decades long period of open ocean migration, remaining within the warmer waters of the Sargasso Sea and the North Atlantic Subtropical Gyre, the circular current system that surrounds it (Figure 2). As adults, they continue to navigate between feeding grounds or back to nesting beaches.

Considering the small size of hatchlings and juveniles, sustained swimming is energetically demanding and there is clear benefit to simply drifting within the convenient conveyor belt of the North Atlantic Gyre. Yet, such simplistic behaviour could come with a risk if the stream branches, such as in the North Atlantic where it splits into separate streams heading south (towards the warmer waters of the Azores) or north (into the colder waters of Ireland and the North Atlantic), Figure 2; drifting into the latter could transport turtles into perilously cold waters. Consequently, it is likely that some degree of positional awareness and navigation is employed and an increasing volume of evidence has emerged on the potential for turtles to follow a magnetic compass [28], exploiting the information provided by the Earth's magnetic field. Such a capacity would clearly be advantageous: despite its diurnal and secular variation, magnetic field information is always available (unlike, say, celestial cues).

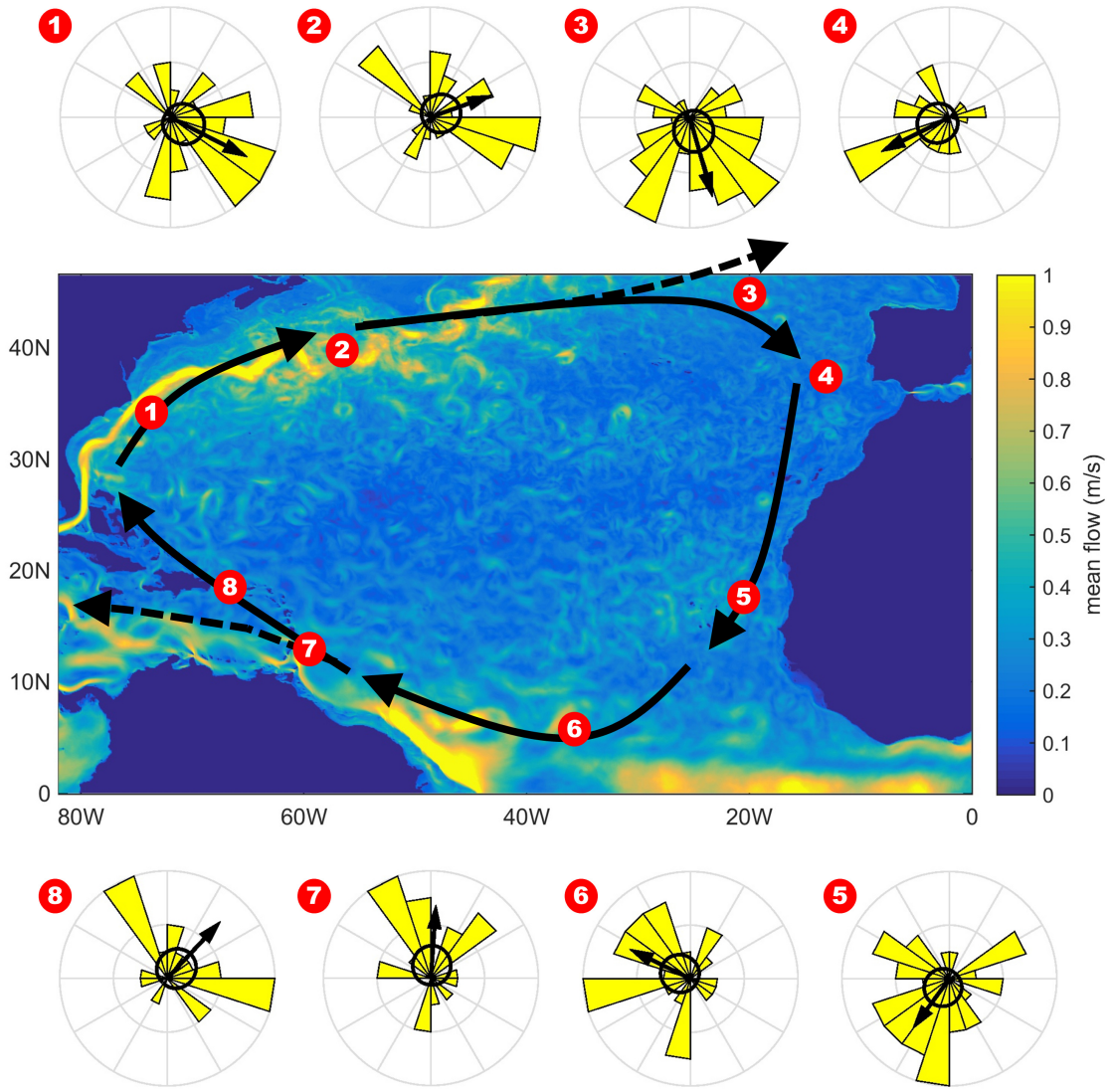


**Fig. 2** The North Atlantic Gyre (Black arrows) is a circular system of currents, formed by the Gulf Stream, the North Atlantic Current, the Canary Current and the North Equatorial Current. For North Atlantic loggerhead turtles, such as those hatching along Florida beaches, remaining inside the region enclosed by the Gyre is optimal for access to suitable feeding grounds (e.g. the Sargasso Sea, the Azores) and to avoid straying into perilously cold waters (e.g. far North Atlantic) or unfamiliar geographic regions (far from traditional nesting/feeding sites). Two potentially hazardous points are indicated by the North Easterly point (3) and the South Westerly point (7): here, currents split into northerly/southerly streams for (3) and northerly/westerly streams for (7). Circular histograms reproduce the hatchling orientation data from [28], where (1-8) correspond to the locations where the magnetic field was reproduced in an experimental arena. When this data is fitted to the von Mises distribution, equation (11), a clear bias emerges, with the dominant direction and concentration strength reflected by the arrow direction and length (concentration parameters $\kappa$ range from 0.67 for dataset 5 to 0.91 for dataset 1). Clearly, the unimodal von Mises distribution may not always be an "optimal" distribution: for example, datasets 2 and 8 may be more convincingly fitted by a multimodal form, such as linear combinations of von Mises distributions. Given the present study aims and the limited sample sizes, we restrict our fitting to the unimodal von Mises distribution.

To investigate this hypothesis, Lohmann and colleagues (see [28] for a review) devised a laboratory experiment that monitors how hatchling orientation changes when exposed to distinct magnetic fields. Briefly, a turtle is placed in a large water-tank while harnessed and tethered to an electronic monitor that computes its swimming direction. The tank is surrounded by a coil system capable of replicating specific geomagnetic fields, such as those found at distinct points along a turtles typical migratory route. Following an acclimatisation period, the mean swimming direction over a 5 minute period is recorded for each turtle, generating orientation data at an individual level. In Figure 2 we reproduce the data summarised in [28] (itself summarising the collection of studies found in [27, 50, 15]). Specifically, magnetic fields were reproduced for different points along the North Atlantic Gyre and, for each location, the (mean) orientation of each tested turtle is binned into a circular histogram. The key inference from these studies is that hatchlings indeed show subtle changes to their preferred swimming direction, consistent with an orientation that optimises remaining within the Gyre. In Section 4.2 we will use this data to parametrise stochastic and continuous models, assessing the capacity for oriented swimming to maintain successful circulation of hatchlings.

### *1.3 Outline*

In the next section (Section 2) we introduce advection-diffusion equations and the fully-anisotropic advection-diffusion framework. We introduce *position-jump* and *velocity-jump* random walks as two alternative stochastic models for oriented movement, and show how these models can be parametrised by translating between individual-level and population-level measurements via circular statistics. In Section 3 we give detailed derivations of the fully-anisotropic advection-diffusion model, starting from either a position-jump or velocity-jump process. In Section 4 we return to the two applications/datasets described above. While each dataset offers a rather distinct set of summary statistics, we show how they can both be incorporated within our framework to parametrise models.

## 2 Basic Tools

Here we outline the basic set of tools that we employ to model and analyse population spread in an anisotropic/oriented environment: advection-diffusion equations, scaling limits for random walks, position-jump and velocity-jump random walks and directional statistics. We note that the derivations of the following sections require a copious notation, spanning scalar, vector and tensor/matrix quantities. To help the reader keep track, we use normal face fonts for scalar quantities (e.g. $t, p, u \ldots$), bold faces for vectors (e.g. $\mathbf{a}, \mathbf{n}, \mathbf{v} \ldots$) and double struck ($\mathbb{D}, \mathbb{V} \ldots$) for tensors and matrices. Much of the material here is of an elementary textbook nature, and we limit references as follows: for more information on the use of advection-diffusion equations in biology, see for example [35, 37]; for more information and perspectives on random walks and their continuous approximations in biological systems, see for example [46, 38, 39, 47, 9, 41, 22]; for more information on the theory and use of directional statistics in biology, see [2, 31].

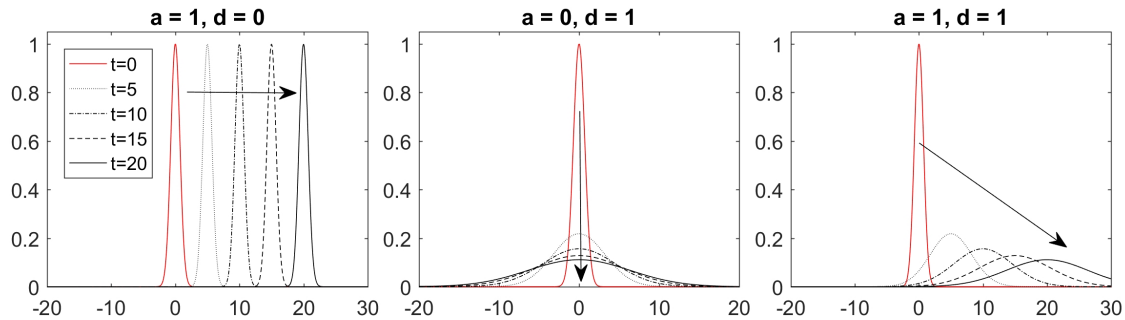


**Fig. 3** Typical solutions of the basic diffusion-advection equation (1). Initial conditions are $u(x,0) = e^{-x^2}$ and solutions shown for (left to right): pure advection; pure diffusion; diffusion-advection.

## *2.1 Advection-Diffusion Equations*

Advection-diffusion equations (AD equations) occupy a prominent position in biological movement modelling [35, 37]. Firstly, AD equations have a relatively straightforward and intuitive form and their long history has generated numerous methods for their analysis. Secondly, AD equations can arise as a limiting form from more realistic/detailed models: they can be derived from discrete and continuous random walks [38], from stochastic differential equations [18] and from individual based models [12]. Thirdly, they have shown to be powerful models capable of describing a wide range of applications in areas as diverse as microbiology [11], ecology [34, 30], physiology [26], and medicine [45]. In short, AD equations describe the basic elements of a movement process.

In the simplest case we restrict to a one-dimensional line and consider a constant drift velocity $a$ and constant diffusion coefficient $d > 0$. The AD equation for some population density $u(x,t)$, where $x$ denotes position along the line and $t$ describes time, is given by

$$u_t + au_x = du_{xx} \tag{1}$$

where the index notation denotes partial derivatives.

In the absence of diffusion ($d = 0$), we have $u_t + au_x = 0$ and solutions are of the form $u(x - at)$, describing movement with constant speed $a$. If $a > 0$ this movement is to the right and if $a < 0$ to the left (see Figure 3 left). In the absence of advection ($a = 0$) we obtain a pure heat (or diffusion) equation $u_t = du_{xx}$: solutions disperse (Figure 3 middle) and (for $x \in \mathbb{R}$) the fundamental solution is

$$u(x,t) = \frac{1}{\sqrt{4\pi dt}} e^{-x^2/4dt}.$$

Taking both terms together ($a \neq 0, d > 0$) the population is transported with velocity $a$ while simultaneously spreading due to diffusion (Figure 3 right).

While the basic elements of directed movement (via $a$) and spatial spread (via $d$) are already contained in (1), questions arise concerning their specific choices related to biological observations/properties: How does

the direction and thickness of nano-grooves translate to advection/diffusion terms? How can we link datasets on turtle headings to these parameters? To answer questions like these we need to generalise the above AD equations (1) in a number of ways:

- advection and diffusion coefficients will more generally depend on space and time;
- we need to explore AD equations in higher space dimensions, in particular two dimensions for the examples studied here;
- as we shall see, any underlying anisotropy or oriented information in the environment can affect both advection and diffusion, necessitating usage of an anisotropic formulation with $n \times n$ diffusion tensor $\mathbb{D}(\mathbf{x},t)$.

Instead of (1) we will therefore consider the *fully anisotropic advection-diffusion equation* (FAAD equation):

$$u_t + \nabla \cdot (\mathbf{a}(\mathbf{x},t)u) = \nabla\nabla : (\mathbb{D}(\mathbf{x},t)u)\,. \tag{2}$$

Note that because the advective velocity ($\mathbf{a}(\mathbf{x},t)$) now depends on space, it appears inside the divergence such that $u_t + \nabla \cdot (\mathbf{a}(\mathbf{x},t)u) = 0$ is a conservation law. The new anisotropic diffusion term in (2) demands special attention. The colon notation (:) used here denotes the contraction of two tensors, and generates a summation across the full suite (i.e. including mixed) of second order derivatives:

$$\nabla\nabla : (\mathbb{D}(\mathbf{x},t)u) = \sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}(\mathbb{D}^{ij}(\mathbf{x},t)u(\mathbf{x},t))\,. \tag{3}$$

Note moreover that this term can be expanded into

$$\nabla\nabla : (\mathbb{D}u) = \nabla \cdot (\mathbb{D}\nabla u) + \nabla \cdot ((\nabla \cdot \mathbb{D})u)\,,$$

which reveals a standard (Fickian-type) anisotropic diffusion term along with an advection term with velocity $\nabla \cdot \mathbb{D}$. As we will show below, the term (3) arises naturally from a detailed random walk description for moving biological agents. We also note that this term can confer some advantages over the standard Fickian anisotropic diffusion form ($\nabla \cdot (\mathbb{D}\nabla u)$): in particular, (3) can allow local maxima and minima to form in the population density steady state distribution, consistent with certain biological observations. Before we move on to this we first show how explicit expressions can be obtained for drift and diffusion terms, correlating to the inputs into an individual-level random walk, and introduce scaling methods in the process.

## *2.2 Scaling Limits for a Simple Random Walk*

Consider an unfortunate hare confined to a life of consecutive and equispaced hops left or right along an infinite one-dimensional road. This animal's convenient movement path can be characterised by a probability density function $p(x,t)$, denoting the probability of the hare being at position $x$ at time $t$. We set $\delta$ to be the hop length, $q$ and $1-q$ as the probabilities of a jump to the right or left and introduce $\tau$ as the (assumed constant) time between consecutive hops. To determine an equation for $p(x,t+\tau)$ we need to calculate the probability of finding the individual at $x$ at time $= t+\tau$. Clearly this will only be possible if the individual has jumped right from position $x-\delta$, or left from $x+\delta$, at time $t$. As a result, we have the discrete *Master equation*

$$p(x,t+\tau) = qp(x-\delta,t) + (1-q)p(x+\delta,t)\,. \tag{4}$$

How can we determine a continuous limit for this discrete equation? The first step is to reinterpret $p$ as a continuous probability distribution and then expand the left hand side about $(x,t)$ as a function of $t$ in powers of $\tau$, and the right hand side terms as functions of $x$ in powers of $\delta$. After removing the arguments $(x,t)$ for clarity, we find

$$p+\tau p_t+\frac{\tau^2}{2}p_{tt}+\ldots=q\left(p-\delta p_x+\frac{\delta^2}{2}p_{xx}-\ldots\right)+(1-q)\left(p+\delta p_x+\frac{\delta^2}{2}p_{xx}+\ldots\right),$$

where the subscripts denote partial derivatives. Simplifying, we obtain

$$p_t(x,t)=\frac{\delta}{\tau}(1-2q)p_x(x,t)+\frac{\delta^2}{2\tau}p_{xx}(x,t)+O(\tau,\frac{\delta^3}{2\tau}). \tag{5}$$

Glancing at Equation (5) hints at the continuous model, where we see that the *leading terms* form an advection-diffusion equation,

$$p_t(x,t)=-ap_x(x,t)+dp_{xx}(x,t) \tag{6}$$

with

$$a=\frac{\delta}{\tau}(2q-1) \quad \text{and} \quad d=\frac{\delta^2}{2\tau}.$$

However, to do this more formally we must think carefully about different *scalings*, corresponding to distinct limiting scenarios as $\delta,\tau\to 0$ and $q\to 1/2$. We will present three choices: others certainly exist, yet the majority do not lead to a useful limit equation. In other words, if $\delta,\tau$ and $q$ do not scale as indicated below, then the above does not provide an appropriate method for deriving a useful continuous model. Note that for each of these scalings, all of the hidden lower order terms of equation (5) limit to zero and are henceforth excluded from consideration.

**(a)** Suppose $\delta,\tau\to 0$ such that $\frac{\delta}{\tau}\to\alpha=$ constant. This describes a *hyperbolic scaling*. Hence, $\frac{\delta^2}{\tau}\to 0$, and the diffusive term vanishes. Thus, we are left with a simple *transport equation*

$$p_t+ap_x=0,$$

where the advective velocity is $a=\alpha(2q-1)$. We can see from this that the advective speed reaches a maximum of $\alpha$ when $q=0$ or 1, which corresponds to always choosing left or always choosing right: i.e. there will be no doubling back.

**(b)** Suppose $\delta,\tau\to 0$ such that $\frac{\delta^2}{\tau}\to 2d=$ constant. This describes a *parabolic scaling*. Here we can consider two cases:

**(b.1)** If $q=\frac{1}{2}$. Here we have $a=0$ and we hence obtain a pure *diffusion equation*

$$p_t=dp_{xx}.$$

**(b.2)** If $q\to\frac{1}{2}$ in such a way that $\frac{\delta}{\tau}(2q-1)\to a$, and $\frac{\delta^2}{2\tau}\to d$, then the scaling results in the *advection-diffusion equation*

$$p_t+ap_x=dp_{xx}. \tag{7}$$

Summarising:

- When $\delta$ and $\tau$ scale in the same way, then we obtain a pure transport equation. This case is called *drift dominated*.

- When $\delta^2 \sim \tau$, we have the *diffusion dominated* case.
- Only if $q - \frac{1}{2} \sim \delta$ do we get both terms, an advection and a diffusion term (*mixed case*). In this case we exactly derive our simple one-dimensional AD equation (1), but now we have a connection from the macroscopic parameters $a$ and $d$ to the statistical inputs of the underlying random walk process $(q, \delta, \tau)$.

The question of which scaling to apply will typically come down to the appropriate relationship between the *macroscopic* and the *individual* spatial and temporal scales: i.e. between the scales of the individual movement process and the scale of the problem. For example, for the hops of a hare their frequency may take place on a timescale of seconds, over a distance of several centimetres. For modelling purposes, we may be interested in the dynamics of the system over observational scales ranging from minutes and metres to years and kilometres. The comparison between these scales provides the key to the appropriate scaling.

It is important to note that we have, in fact, only derived a continuous limiting equation for the probability distribution of finding an individual at position $x$ at time $t$. Can we directly relate $p$ to a density function $u$ that describes the distribution of a population? Formally, this would require that the jumpers are stochastically independent, i.e. that any interactions between population members can be (reasonably) ignored. This would, quite obviously, be a strong assumption if applied generally and its validity demands careful assessment [46, 54]. Accounting for population interactions will significantly complicate the proceedings (often to the point of intractability) and we shall therefore restrict to stochastically independent jumpers in the context of this chapter: effectively, we directly interchange the probability distribution $p$ with the population density distribution $u$.

## 2.3 Classes of Biological Random Walks

In the above example we considered an *uncorrelated position-jump random walk on a discrete and regular one-dimensional lattice* for our underlying movement process: moves were uncorrelated, in that the decision of which direction to take did not depend on the previous decision(s), movement occurred through positional jumps in space that ignored explicit description of passage between successive points, and were of fixed length, so that the path was localised to equally-spaced points along a one-dimensional line.

More generally, two popular random walk descriptions for biological movement are the *position-jump* and *velocity-jump* random walk processes. These descriptions have been introduced to biological modelling by Othmer, Dunbar and Alt [38] and subsequently proven to be powerful and popular approaches. In the simpler position-jump process, the random walker jumps discretely from point to point according to certain jump probabilities (Figure 4 left); the one-dimensional random walk discussed above provides a particularly simple example. The more sophisticated velocity-jump process assumes piecewise continuous movement through space, with random walkers changing their velocity (or heading) during *turns*. Choosing an appropriate random walk description involves a balancing of their respective advantages: for example, while the velocity-jump approach benefits from its more natural representation of biological movement, the subsequent derivation of a continuous limiting equation is somewhat more complicated (Figure 4 right).

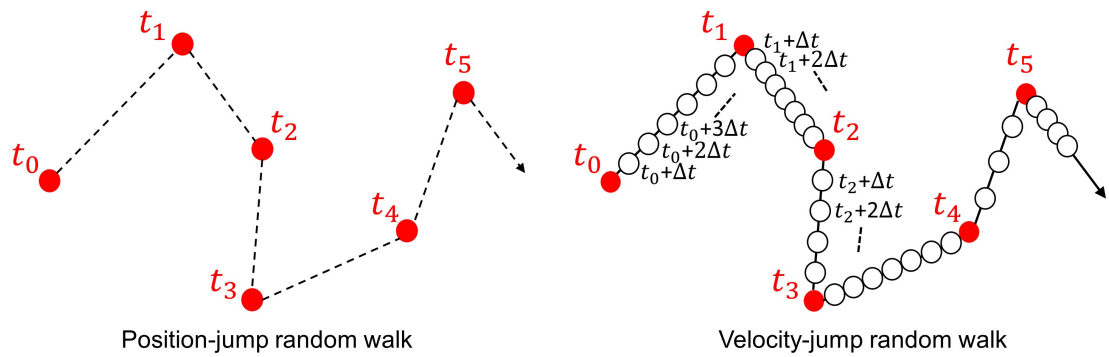


**Fig. 4** Schematic illustrating position-jump and velocity-jump random walks. (Left) In the position-jump process, the particle makes instantaneous jumps through space at discrete times $t_0, t_1, t_2, \ldots$. (Right) In the velocity-jump process, the particle makes instantaneous velocity-changes at discrete times $t_0, t_1, t_2, \ldots$ (red circles), but subsequently moves continuously through space with a fixed velocity in the intervening times (white circles).

### 2.3.1 Position-Jump Processes

Moving beyond our simple random walk above, a more general position-jump random walk assumes movement proceeds through a sequence of positional jumps in space, interspersed according to some characteristic mean waiting time. Such instantaneous transitions are clearly somewhat unrealistic in the context of biological movement, yet given the discrete nature of many datasets (for example, satellite tracking of an animal in which its path is recorded through its spatial coordinate at discrete times) a position-jump model can often be justified as a reasonable approximation [5, 57].

Position-jump random walks can be alternatively stated via a discrete or continuous time master equation [38], and here we consider the former form. Specifically, we consider a population of stochastically independent jumpers performing a discrete time random walk, starting at $t = 0$ and making jumps at fixed times separated by time step $\tau$. We introduce a redistribution kernel $K(\mathbf{y}, \mathbf{x}, t)$, a probability density function for a jump from position $\mathbf{x}$ to $\mathbf{y}$ at time $t$. Note that, as a probability, we have $K \geq 0$.

The difference in the population density at $\mathbf{x}$ between times $t$ and $t + \tau$ will be determined by summing all jumps into position $\mathbf{x}$ and subtracting all those away from position $\mathbf{x}$, i.e. by the equation

$$u(\mathbf{x}, t+\tau) - u(\mathbf{x}, t) = \int_{D^{\mathbf{x}}} K(\mathbf{x}, \mathbf{y}, t) u(\mathbf{y}, t) - K(\mathbf{y}, \mathbf{x}, t) u(\mathbf{x}, t) d\mu(\mathbf{y}) \,. \tag{8}$$

In the above, $(D^{\mathbf{x}}, \mu(\mathbf{y}))$ is a measure space. The above is general for random walks including jumps of various step lengths, or cases where movement occurs in continuous space or is restricted to discrete jumps between regularly or irregularly arranged nodes. The set $D^{\mathbf{x}}$ determines the set of destination/incoming sites for position $\mathbf{x}$, i.e. the set of points $\mathbf{y} \in D^{\mathbf{x}}$ from which jumps into or out of $\mathbf{x}$ can be made, with $\mu(\mathbf{y})$ its associated measure. For example, if jumps can be made in any direction and any distance up to length $h$, then $D^{\mathbf{x}}$ becomes the ball centred on $\mathbf{x}$ of radius $h$ and the associated measure is the standard Lebesgue measure. If jumps can be made in any direction, but are restricted to a fixed length $h$, then $D^{\mathbf{x}}$ will be the sphere of radius $h$ centred on $\mathbf{x}$ and the associated measure is the surface Lebesgue measure. When movements become

restricted to a set of nodes, $D^{\mathbf{x}}$ becomes a finite or infinite set of discrete positions with a corresponding discrete measure.

The choice of redistribution kernel $K$ is a key modelling decision, and allows various potential factors to be incorporated: for example, $K$ could incorporate an impact due to environmental anisotropy or navigating cues that bias jumps into particular headings. The redistribution kernel is taken to be a probability measure, i.e.

$$\int_{D^{\mathbf{x}}} K(\mathbf{y},\mathbf{x},t)d\mu(y) = 1\,.$$

The above excludes spatio-temporal variation in the rate that jumps are made. However, it is noted that this is distinct from variation in staying at the same site, since $D^{\mathbf{x}}$ could include $\mathbf{x}$ and remaining would correspond to $K(\mathbf{x},\mathbf{x},t) > 0$.

### 2.3.2 Velocity-Jump Processes

In velocity-jump random walks, movement consists of smooth runs with constant velocity interspersed by (instantaneous) reorientations [38]. For stochastically independent walkers, the individual-scale velocity-jump random walk can be formulated as an individual-scale continuous *transport equation*. Transport models form a powerful and relatively new tool in the modelling and analysis of animal and cell movement [37, 21, 40, 47], although they have a long history in continuum mechanics (where they are usually referred to as kinetic equations) [8, 3]. As a result, various tools and techniques have been developed and in particular the scaling techniques that allow their approximation to a reduced (and hopefully simpler) macroscopic model [47, 22]. Consequently, the transport equation can be thought of as a bridge that connects the individual random walk to a fully continuous macroscopic model.

The reapplication of transport equations to biological processes has grown from seminal work of the 1980s (see [1, 38]) as an approach for modelling biological movement, whether by cells or organisms. Transport equations typically refer to mathematical models in which the particles of interest are structured by their position in space, time and velocity. In words, the transport equation for animal/cell movement takes the intuitively simple form:

$$\begin{array}{ccc} \text{Rate of change of population} & \text{Change due to} & \text{Change due to} \\ \text{moving with velocity } \mathbf{v} & = \text{movement through} + & \text{turning into or out} \\ \text{at position } \mathbf{x} \text{ time } t & \text{space} & \text{of velocity } \mathbf{v} \end{array}$$

Formally, if we define by $p(\mathbf{v},\mathbf{x},t)$ to be the density of the population moving with velocity $\mathbf{v} \in V$ at position $\mathbf{x}$ and time $t$, then

$$p_t(\mathbf{v},\mathbf{x},t) + \mathbf{v}\cdot\nabla p(\mathbf{v},\mathbf{x},t) = \mathscr{L}p(\mathbf{v},\mathbf{x},t)\,, \tag{9}$$

where $\mathscr{L}$ denotes a *turning operator* that describes the process of velocity switching[1]. For the velocity space $V \subset \mathbb{R}^n$ we take $V = [s_1, s_2] \times S^{n-1}$, where $0 \le s_1 \le s_2 < \infty$, $s_1$ and $s_2$ define the lower and upper bounds for organism movement speed[2] and $S^{n-1}$ defines the unit sphere.

[1] We note that this particular form assumes there is no net force on the particles, and thus no inertia on them.

[2] It is worth noting that this is a key distinction from the kinetic theory of gas molecules, where $V = \mathbb{R}^n$ permits (at least theoretically) individual molecules to acquire infinite momentum [8].

The choice of $\mathscr{L}$ forms a key modelling decision, and an oft-used form is the integral operator representation [38]:

$$\mathscr{L}p(\mathbf{v},\mathbf{x},t) = -\mu p(\mathbf{v},\mathbf{x},t) + \mu \int_V T(\mathbf{v},\mathbf{v}',\mathbf{x},t)p(\mathbf{v},\mathbf{x},t)d\mathbf{v}', \tag{10}$$

where the first term on the right hand side gives the rate at which particles switch away from velocity $\mathbf{v}$ and the second term denotes the switching into velocity $\mathbf{v}$ from all other velocities. The parameter $\mu$ is the *turning rate*, with $1/\mu$ the *mean run time* between individual turns. The turning kernel $T(\mathbf{v},\mathbf{v}',\mathbf{x},t) \geq 0$ denotes the switching into velocity $\mathbf{v}$ for a turn made at position $\mathbf{x}$ and time $t$, given some previous velocity $\mathbf{v}'$. Mass conservation demands

$$\int_V T(\mathbf{v},\mathbf{v}',\mathbf{x},t)d\mathbf{v} = 1$$

and consequently $T$ denotes a probability measure over $V$. As for the redistribution kernel in the position-jump process, its choice is a major consideration: for example orientation signals from the environment at $\mathbf{x}$ and time $t$, or the inclusion of persistence in the previous direction $\mathbf{v}'$.

## *2.4 Directional Statistics*

Each of the position-jump and velocity-jump processes above rely on various biological inputs: mean waiting times, speeds, turning rates and redistribution kernels. It is through these inputs that the random walk can be linked to biological datasets, and not least significant are the kernels $K$ and $T$, which respectively describe probability distribution functions for either the redistribution kernel for a positional jump from some position $\mathbf{x}$ to a position $\mathbf{y}$, or a change of velocity from $\mathbf{v}'$ to $\mathbf{v}$. Fundamentally, each distribution encapsulates an orientating "choice" of the animal or cell and we now turn to consider some suitable representations.

Typical datasets for cell movement and animal navigation problems relate to orientations/headings in space and handling such data demands a review of some concepts from directional statistics [31]. In two dimensions, directional (or circular) statistics involves consideration of data on orientations that can be expressed with respect to some angle $\alpha$ relative to a given $x$-direction. The problem of directly transposing the definitions of regular (linear) statistics to circular statistics becomes immediately apparent with even its simplest concepts: for a set of angles uniformly distributed across the circle, what meaning would the (linear) mean angle of this dataset have?

In general we consider the set of directions on the $n$-dimensional sphere, i.e. the set of unit vectors $\mathbf{n} \in S^{n-1}$. A directional distribution is then a probability distribution $q(\mathbf{n})$ defined over $S^{n-1}$, i.e. one satisfying

$$q(\mathbf{n}) \geq 0 \quad \text{and} \quad \int_{S^{n-1}} q(\mathbf{n})d\mathbf{n} = 1\,.$$

Of particular importance for our work are the first and second moments of $q$, respectively the expectation $\mathbf{E}_q$ and variance-covariance matrix $\mathbb{V}_q$ (which we will often refer to simply as the variance):

$$\begin{aligned}\mathbf{E}_q &= \int_{S^{n-1}} \mathbf{n}q(\mathbf{n})d\mathbf{n},\\ \mathbb{V}_q &= \int_{S^{n-1}} (\mathbf{n}-\mathbf{E}_q)(\mathbf{n}-\mathbf{E}_q)^T q(\mathbf{n})d\mathbf{n}.\end{aligned}$$

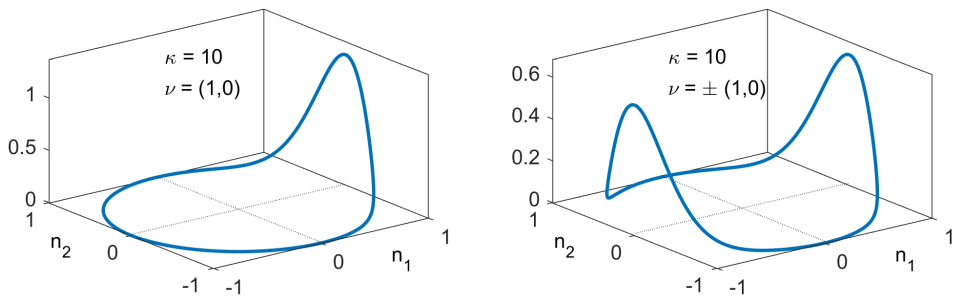


**Fig. 5** Left: The unimodal von Mises distribution as a function of $\mathbf{n} = (n_1, n_2)^T \in S^1$ with a peak at $\boldsymbol{\nu} = (1,0)^T$. Right: The bimodal von Mises distribution $q_{vM}$ as a function of $\mathbf{n} \in S^1$ with peaks at $\boldsymbol{\nu} = \pm(1,0)^T$. In these plots we set $\kappa = 10$.

In two dimensions, distributions will be defined on the unit circle, i.e. $\mathbf{n} \in S^1$. The simplest example is the uniform distribution, $q(\mathbf{n}) = \frac{1}{2\pi}$, although this has obviously limited usage in cases where data shows clear clustering/structure.

Given the enormous importance of the normal distribution in linear statistics, it is clearly desirable to define a similar concept for circular statistics. While the wrapped normal distribution offers the most direct analogue, the normal distribution's prominent position in circular statistics is filled instead by its sibling the *von Mises distribution* [31, 2], which benefits from its more analytically tractable form; the subtle differences between the wrapped normal and von Mises distribution are unlikely to be differentiated within the context of typical (noisy) biological datasets. Suppose we have some dominant/preferred direction $\boldsymbol{\nu} \in S^1$, then the von Mises distribution is given by

$$q_{vM}(\mathbf{n}, \boldsymbol{\nu}, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \mathbf{n} \cdot \boldsymbol{\nu}} \tag{11}$$

for $\mathbf{n} \in S^1$. Here $\kappa$ denotes the *concentration parameter* and $I_0(\kappa)$ ($I_j(\kappa)$) denotes the modified Bessel function of first kind of order 0 (order $j$). The von Mises distribution is illustrated in Figure 5 on the left.

It is, of course, equally possible to write down the von Mises distribution in terms of polar angles. Denoting $\alpha$ to be the angle of $\mathbf{n}$ and $\phi$ to be the angle of $\boldsymbol{\nu}$ (i.e. the *dominant angle*), then we can write

$$q_{vM}(\alpha, \phi, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\alpha - \phi)} .$$

The above form is more common, particularly in the biological literature [32], but it is less useful for computations and can be notationally more cumbersome. Hence we work with the coordinate free form (11) when possible.

As for the normal distribution on the line, the von Mises distribution on the circle is the workhorse of planar directional statistics [31, 2]. It can be derived from random walks, diffusion equations and energy principles, and has applications in earth sciences, physics, biology, medicine and elsewhere. It is used for data fitting and hypothesis testing of directional data, and we will use it here for our modelling of biological movement. The first and second moments of (11) have been computed in [23] (amongst elsewhere) and are given by

$$\mathbf{E}_{q_{vM}} = \frac{I_1(\kappa)}{I_0(\kappa)}\,\boldsymbol{\nu}\,; \tag{12}$$

$$\mathbb{V}_{q_{vM}} = \frac{1}{2}\left(1-\frac{I_2(\kappa)}{I_0(\kappa)}\right)\mathbb{I}_2 + \left(\frac{I_2(\kappa)}{I_0(\kappa)} - \left(\frac{I_1(\kappa)}{I_0(\kappa)}\right)^2\right)\boldsymbol{\nu}\boldsymbol{\nu}^T\,. \tag{13}$$

Note that $\mathbb{I}_2$ denotes the $2\times 2$ identity matrix, and $\boldsymbol{\nu}\boldsymbol{\nu}^T$ denotes the dyadic product of two vectors (in tensor notation $\boldsymbol{\nu}\otimes\boldsymbol{\nu}$).

Many biological datasets possess multimodal structure and we note that the von Mises distrubution can be extended to describe such instances, for example through simple linear combinations of (11); the moments correspondingly follow from linear combinations of (12-13). A particularly useful case emerges for axially-symmetric directional information, such as the spreading of cells along nanogrooves or animal movement along linear environment structures such as seismic lines [33]. In such cases we can define a bimodal von Mises distributions with equal sized local maxima at $\pm\boldsymbol{\nu}$. As shown in [23], we find that for given $\boldsymbol{\nu}\in S^1$ the bimodal von Mises distribution

$$q_{bvM}(\mathbf{n},\boldsymbol{\nu},\kappa) = \frac{1}{4\pi I_0(\kappa)}\left(e^{\kappa\mathbf{n}\cdot\boldsymbol{\nu}} + e^{-\kappa\mathbf{n}\cdot\boldsymbol{\nu}}\right), \tag{14}$$

has moments

$$\mathbf{E}_{q_{bvM}} = \mathbf{0}\,, \tag{15}$$

$$\mathbb{V}_{q_{bvM}} = \frac{1}{2}\left(1-\frac{I_2(\kappa)}{I_0(\kappa)}\right)\mathbb{I}_2 + \frac{I_2(\kappa)}{I_0(\kappa)}\boldsymbol{\nu}\boldsymbol{\nu}^T. \tag{16}$$

An illustration of the bimodal von Mises distribution is shown in Figure 5 on the right.

For the present chapter we exclusively concentrate on two-dimensional applications, however it is worth remarking that extensions can be made to three dimensions. The equivalent of the von Mises distribution in three dimensions is called the *Fisher distribution* and is given by

$$q_F(\mathbf{n},\boldsymbol{\nu},\kappa) = \frac{\kappa}{4\pi\sinh(\kappa)}e^{\kappa\mathbf{n}\cdot\boldsymbol{\nu}}, \qquad \mathbf{n}\in S^2. \tag{17}$$

Again, first and second moments have been previously calculated for this distribution (see [23]), given by

$$\mathbf{E}_{q_F} = \left(\coth\kappa - \frac{1}{\kappa}\right)\boldsymbol{\nu}\,, \tag{18}$$

$$\mathbb{V}_{q_F} = \left(\frac{\coth\kappa}{\kappa} - \frac{1}{\kappa^2}\right)\mathbb{I} + \left(1 - \frac{\coth\kappa}{\kappa} + \frac{2}{\kappa^2} - \coth^2\kappa\right)\boldsymbol{\nu}\boldsymbol{\nu}^T. \tag{19}$$

## 3 Derivation of Fully Anisotropic Advection-Diffusion Equations

Here we present two derivations of the FAAD model (2), respectively from a position-jump and velocity-jump process. We will find that both the macroscopic drift velocity $\mathbf{a}$ and the diffusion tensor $\mathbb{D}$ depend

on statistical properties of the parameters in the corresponding random walk model. Hence, the choice of an appropriate model can be linked to the available data: if we can compute mean and variance of species locations, then the position-jump framework applies (see our cell movement example); if the data allow estimates for mean speeds, mean directions and their variances, then the velocity-jump process is perhaps a better choice (see the sea-turtle example).

### *3.1 Position-Jump Derivation*

For the position-jump derivation we will make a number of convenient restrictions:

1. we assume random walks in which the jumps can occur in any direction (i.e. lattice-free), but are restricted to fixed length $\delta$.
2. we assume the jump is *myopic* (or short-sighted).

The first restriction determines that the set $D$ in equation (8) simply becomes the sphere of radius $\delta$. The myopic nature of the jump implies that the heading is based only on environmental information obtained at the present site, i.e. at $(\mathbf{x},t)$ for a walker at position $\mathbf{x}$ at time $t$; alternatives could involve, as an example, a dependence on information at the destination site, or a comparison between the current and destination site [55].

The consequence of these assumptions is that our redistribution kernels can be written in terms of a directional distribution for choosing direction $\mathbf{n} \in S^{n-1}$, i.e. $K(\mathbf{y},\mathbf{x},t) = k(\mathbf{n},\mathbf{x},t)$ where $\mathbf{n}$ is in the direction $\frac{\mathbf{y}-\mathbf{x}}{|\mathbf{y}-\mathbf{x}|}$ and the Master equation becomes

$$u(\mathbf{x},t+\tau) - u(\mathbf{x},t) = \int_{S^{n-1}} k(\mathbf{n},\mathbf{x}-\delta\mathbf{n},t)u(\mathbf{x}-\delta\mathbf{n},t) - k(\mathbf{n},\mathbf{x},t)u(\mathbf{x},t)d\mathbf{n}\,. \tag{20}$$

At this point it is interesting to quickly consider the connection to the one-dimensional case (4) that was studied earlier. In the one-dimensional case we have only two headings, $\mathbf{n} \in \{-1,1\}$. Hence we define

$$k(\mathbf{n},x,t) = q\delta_0(-1-\mathbf{n}) + (1-q)\delta_0(1-\mathbf{n}),$$

where $\delta_0$ denotes the Dirac-delta distribution. Then (20) becomes

$$u(x,t+\tau) = qu(x-\delta,t) + (1-q)u(x+\delta,t),$$

which is exactly (4).

For small values of $\delta$ and $\tau$ we expand the right hand side of equation (20) about $\mathbf{x}$ and the left hand side about $t$ to obtain

$$\begin{aligned}\frac{\partial u}{\partial t} + O(\tau) &= \frac{\delta}{\tau}\int_{S^{n-1}} -\mathbf{n}\cdot\nabla(ku) + \frac{\delta}{2}(\mathbf{n}\cdot\nabla)^2(ku) + O(\delta^2)d\mathbf{n}\,,\\ &= -\frac{\delta}{\tau}\left(\nabla\cdot\int_{S^{n-1}} k\mathbf{n}d\mathbf{n}\right)u + \frac{\delta^2}{2\tau}\left(\nabla\nabla : \int_{S_{n-1}} \mathbf{n}\mathbf{n}^T k d\mathbf{n}\right)u + O(\delta^3/\tau)\,,\end{aligned}$$

where we use the colon notation (:) which denotes the contraction of two tensors as

$$A : B = \sum_{i,j=1}^{n} a_{ij} b_{ij}, \qquad A, B \in \mathbb{R}^{n\times n}.$$

As discussed in Section 2.2, distinct scalings generate different continuous limits and we again consider both the drift and diffusion dominated scenarios.

- (drift dominated) if $\delta, \tau \to 0$ such that $\lim_{\delta,\tau\to 0} \frac{\delta}{\tau} = c$ (constant) we have the hyperbolic model

$$\frac{\partial u}{\partial t} + \nabla \cdot (\mathbf{a}(\mathbf{x},t)u) = 0\,,$$

  where $\mathbf{a}(\mathbf{x},t) = c \int_{S^{n-1}} \mathbf{n} k(\mathbf{n},\mathbf{x},t) d\mathbf{n}$ (i.e. the advection is proportional to the first moment of $k$).
- if $\delta, \tau \to 0$ such that $\lim_{\delta,\tau\to 0} \frac{\delta^2}{2\tau} = d$ then we have two cases
  - (diffusion dominated) if $\int_{S^{n-1}} \mathbf{n} k d\mathbf{n} = 0$ then we have

$$\frac{\partial u}{\partial t} = \nabla\nabla : (\mathbb{D}(\mathbf{x},t)u)\,,$$

    where $\mathbb{D}(\mathbf{x},t)$ is the $n \times n$ matrix defined by $\mathbb{D}(\mathbf{x},t) = d \int_{S^{n-1}} \mathbf{n}\mathbf{n}^T k(\mathbf{n},x,t) d\mathbf{n}$.
  - (drift-diffusion) If $\lim_{\delta,\tau\to 0} \frac{\delta^2}{2\tau} = d$ and $\lim_{\delta,\tau\to 0} \frac{\delta}{\tau} \int_{S^{n-1}} \mathbf{n} k d\mathbf{n} \sim c\delta$ we have

$$\frac{\partial u}{\partial t} + \nabla \cdot (\mathbf{a}(\mathbf{x},t)u) = \nabla\nabla : (\mathbb{D}(\mathbf{x},t)u)\,,$$

    with

$$\begin{aligned}
\mathbf{a}(\mathbf{x},t) &= c \int_{S^{n-1}} \mathbf{n} k(\mathbf{n},\mathbf{x},t) d\mathbf{n}\,,\\
\mathbb{D}(\mathbf{x},t) &= d \int_{S^{n-1}} (\mathbf{n} - \mathbf{a}(\mathbf{x},t))(\mathbf{n} - \mathbf{a}(\mathbf{x},t))^T k(\mathbf{n},\mathbf{x},t) d\mathbf{n}\,.
\end{aligned}$$

The final form is particularly relevant, as it is exactly the FAAD model we introduced earlier. In this case, we now have a connection to the advection velocity and diffusion tensor terms from the underlying statistical inputs $k(\mathbf{n},x,t)$ of a random walk process.

## *3.2 Velocity-Jump Derivation*

To facilitate the derivation we consider a simplified form of transport equation. Specifically, we assume that the turning kernel does not depend on the previous velocity $\mathbf{v}'$, i.e.

$$T(\mathbf{v},\mathbf{v}',\mathbf{x},t) = T(\mathbf{v},\mathbf{x},t)\,.$$

Using this choice in (10) for (9) we have the considerably simpler form

$$p_t(\mathbf{v},\mathbf{x},t) + \mathbf{v} \cdot \nabla p(\mathbf{v},\mathbf{x},t) = -\mu p(\mathbf{v},\mathbf{x},t) + T(\mathbf{v},\mathbf{x},t) u(\mathbf{x},t)\,, \tag{21}$$

where we have defined the macroscopic density

$$u(\mathbf{x},t) = \int_V p(\mathbf{v},\mathbf{x},t)d\mathbf{v}. \tag{22}$$

The process from here is to derive an evolution equation for the macroscopic density $u(\mathbf{x},t)$, which can be achieved through a variety of scaling techniques, including parabolic scaling, hyperbolic scaling and moment closure. For a detailed treatment for model (21) we refer to our earlier paper [22] and we summarise one such choice here: moment closure.

### 3.2.1 Moment Closure Method

In a moment closure approach, the idea is to identify statistically meaningful quantities related to $p$ and $T$, such as expectations and variances. We remind ourselves that the formulation demands that the turning distribution $T(\mathbf{v},\mathbf{x},t)$ is a probability measure, i.e.

$$T(\mathbf{v},\mathbf{x},t) \geq 0, \qquad \int_V T(\mathbf{v},\mathbf{x},t)d\mathbf{v} = 1,$$

and we consider its expectation $\mathbf{E}_T$ and variance $\mathbb{V}_T$ ,

$$\mathbf{E}_T(\mathbf{x},t) = \int_V \mathbf{v}T(\mathbf{v},\mathbf{x},t)d\mathbf{v}, \qquad \mathbb{V}_T(\mathbf{x},t) := \int_V (\mathbf{v}-\mathbf{E}_T(\mathbf{x},t))(\mathbf{v}-\mathbf{E}_T(\mathbf{x},t))^T T(\mathbf{v},\mathbf{x},t)d\mathbf{v}\,. \tag{23}$$

$\mathbf{E}_T(\mathbf{x},t)$ describes the mean new velocity vector for the turning kernel, while $\mathbb{V}_T(\mathbf{x},t)$ is its variance-covariance matrix.

We now introduce the same quantities for $p(\mathbf{v},\mathbf{x},t)$, although we note that $p$ in itself is not a probability measure, since $\int_V p(\mathbf{v},\mathbf{x},t)d\mathbf{v} = u(\mathbf{x},t)$ is not necessarily equal to one. But we can normalise, introducing $\hat{p}$ via the equation

$$u(\mathbf{x},t)\hat{p}(\mathbf{v},\mathbf{x},t) = p(\mathbf{v},\mathbf{x},t)$$

and noting that $\int_V \hat{p}(\mathbf{v},\mathbf{x},t)d\mathbf{v} = 1$. We subsequently introduce the expectation and variances

$$\begin{aligned}\mathbf{E}_{\hat{p}}(\mathbf{x},t) &= \int_V \mathbf{v}\hat{p}(\mathbf{v},\mathbf{x},t)d\mathbf{v},\\ \mathbb{V}_{\hat{p}}(\mathbf{x},t) &= \int_V (\mathbf{v}-\mathbf{E}_{\hat{p}}(\mathbf{x},t))(\mathbf{v}-\mathbf{E}_{\hat{p}}(\mathbf{x},t))^T \hat{p}(\mathbf{v},\mathbf{x},t)d\mathbf{v}.\end{aligned}$$

Then, $\mathbf{E}_{\hat{p}}$ defines the mean velocity of the normalized population while $\mathbb{V}_{\hat{p}}$ is its variance-covariance matrix. In terms of the original population density $p$, we can write

$$\int_V \mathbf{v}p(\mathbf{v},\mathbf{x},t)d\mathbf{v} = \mathbf{E}_{\hat{p}}(\mathbf{x},t)u(\mathbf{x},t)\,, \tag{24}$$

$$\int_V (\mathbf{v}-\mathbf{E}_{\hat{p}}(\mathbf{x},t))(\mathbf{v}-\mathbf{E}_{\hat{p}}(\mathbf{x},t))^T p(\mathbf{v},\mathbf{x},t)d\mathbf{v} = \mathbb{V}_{\hat{p}}(\mathbf{x},t)u(\mathbf{x},t)\,. \tag{25}$$

Next we explain the moment closure method itself. We can derive equations for the expectation and variance introduced above, and it turns out that the equation for the expectation (first moment) depends on the variance (second moment) while the equation for the variance depends on a third moment etc. Effectively we obtain an infinite hierarchy of moment equations, where each new equation depends on a next higher moment. To

obtain a usable model, the sequence of equations must be cut somewhere, a process termed *moment closure*. Generally, choosing the right closure condition is a work of art and many plausible approaches are available in the literature [8, 20]. Here we will choose a standard method that uses the equilibrium distribution and cut at the second moment to obtain a single equation of type (2) for the mass density $u(\mathbf{x},t)$.

Let us start by integrating equation (21) over $V$ and express each term with respect to the corresponding moments. Note that hereon we omit the arguments for readability.

$$\int_V p_t d\mathbf{v} + \int_V \nabla\cdot\mathbf{v}p\, d\mathbf{v} = -\mu\int_V p\, d\mathbf{v} + \mu\int_V T\, d\mathbf{v}\, u,$$

which can equivalently be written as

$$u_t + \nabla\cdot(\mathbf{E}_{\hat{p}}u) = -\mu u + \mu u = 0.$$

Hence our first equation is a conservation law

$$u_t + \nabla\cdot(\mathbf{E}_{\hat{p}}u) = 0\,. \tag{26}$$

As a next step we multiply (21) by $\mathbf{v}$ and again integrate over $V$. We obtain

$$\int_V \mathbf{v}u_t d\mathbf{v} + \int \mathbf{v}(\nabla\cdot\mathbf{v}p)d\mathbf{v} = -\mu\int_V \mathbf{v}p\, d\mathbf{v} + \mu\int_V \mathbf{v}T\, d\mathbf{v}\, u\,,$$

which can be equivalently written as

$$(\mathbf{E}_{\hat{p}}u)_t + \nabla\cdot\int_V \mathbf{v}\mathbf{v}^T p\, d\mathbf{v} = \mu(\mathbf{E}_T - \mathbf{E}_{\hat{p}})u\,. \tag{27}$$

We write the second moment $\int \mathbf{v}\mathbf{v}^T p d\mathbf{v}$ in terms of the variance of $\hat{p}$, i.e.

$$\begin{aligned}\mathbb{V}_{\hat{p}}u &= \int_V (\mathbf{v}-\mathbf{E}_{\hat{p}})(\mathbf{v}-\mathbf{E}_{\hat{p}})^T p d\mathbf{v}\,,\\ &= \int_V \mathbf{v}\mathbf{v}^T p d\mathbf{v} - 2\int_V \mathbf{v}\mathbf{E}_{\hat{p}}^T p d\mathbf{v} + \mathbf{E}_{\hat{p}}\mathbf{E}_{\hat{p}}^T u.\end{aligned}$$

Hence

$$\int_V \mathbf{v}\mathbf{v}^T p d\mathbf{v} = \mathbb{V}_{\hat{p}}u + \mathbf{E}_{\hat{p}}\mathbf{E}_{\hat{p}}^T u.$$

We use this expression in (27) and obtain the equation for the expectation:

$$(\mathbf{E}_{\hat{p}}u)_t + \nabla\cdot(\mathbf{E}_{\hat{p}}\mathbf{E}_{\hat{p}}^T u) = -\nabla\cdot(\mathbb{V}_{\hat{p}}u) + \mu(\mathbf{E}_T - \mathbf{E}_{\hat{p}})u\,. \tag{28}$$

So far we have simply integrated and introduced a few fancy variables for $\mathbf{E}_{\hat{p}}, \mathbb{V}_{\hat{p}}$ etc. The next step is to present two critical assumptions that allow us to close the system:

**(a1) Moment closure –** the variance $\mathbb{V}_{\hat{p}}$ is computed from the equilibrium distribution $p_e$: $\mathbb{V}_{\hat{p}} \approx \mathbb{V}_{\hat{p}_e}$.

**(a2) Fast flux relaxation –** the equation (28) for the expectation $\mathbf{E}_{\hat{p}}$ is in quasi-equilibrium.

It is noted that the above assumptions were originally conceived in a physical context, namely the kinetic theory of dilute gases [8]. The extent to which these can be directly translated to biological particles, such as cells and organisms, is uncertain and a goal for further investigations: within the present article we simply take them as stated. The first assumption has proven to be useful in a number of studies. The second assumption effectively stipulates that, *at the space/time scales* of the macroscopic model, the particle instantaneously respond to local information: reasonable, say, for an organism switching direction multiple times a day but studied over a macroscopic scale of months to years.

The equilibrium distribution $p_e$ can be computed from the condition $\mathscr{L} p_e = 0$ where $\mathscr{L}$ is the integral operator from (10). In our case

$$\mathscr{L} p = \mu(Tu - p) = 0$$

is solved by the equilibrium distribution ,

$$p_e(\mathbf{v},\mathbf{x},t) = u(\mathbf{x},t)T(\mathbf{v},\mathbf{x},t).$$

This equilibrium distribution has the expectation

$$\mathbf{E}_{\hat{p}_e} u = \int_V \mathbf{v} p_e d\mathbf{v} = \int_V \mathbf{v} uT \, d\mathbf{v} = \mathbf{E}_T u\,. \tag{29}$$

Now we approximate the highest order term, the variance as

$$\mathbb{V}_{\hat{p}} \approx \mathbb{V}_{\hat{p}_e} = \int_V (\mathbf{v} - \mathbf{E}_{\hat{p}_e})(\mathbf{v} - \mathbf{E}_{\hat{p}_e})^T uT \, d\mathbf{v} = \mathbb{V}_T u\,. \tag{30}$$

In assumption (a2) we postulate that the equation (28) is in quasi steady state, i.e.

$$\mathbf{0} \approx -\nabla \cdot (\mathbb{V}_{\hat{p}} u) + \mu(\mathbf{E}_T - \mathbf{E}_{\hat{p}})u\,,$$

and, substituting the moment closure (30), we find the approximation

$$\mathbf{E}_{\hat{p}} u \approx -\frac{1}{\mu}\nabla \cdot (\mathbb{V}_T u) + \mathbf{E}_T u\,. \tag{31}$$

Finally, we substitute (31) into the conservation law (26) and we assume that the approximation is good (i.e. we replace $\approx$ with $=$) to obtain a closed system

$$u_t + \nabla \cdot (\mathbf{E}_T u) = \frac{1}{\mu}\nabla\nabla : (\mathbb{V}_T u)\,. \tag{32}$$

This closed equation is exactly the fully anisotropic advection-diffusion equation (FAAD) in (2) with

$$\mathbf{a}(\mathbf{x},t) = \mathbf{E}_T(\mathbf{x},t) \quad \text{and} \quad \mathbb{D}(\mathbf{x},t) = \frac{1}{\mu}\mathbb{V}_T(\mathbf{x},t). \tag{33}$$

Let us consider two special cases of this derivation.

**Example 1: (directional distributions)** Some further simplifications can be used to relate turning directly to a directional distribution. Let us restrict movement to a single speed, i.e. $V = sS^{n-1}$, where $s$ is the mean speed and $S^{n-1}$ is the $n$-dimensional sphere. Hence, $\mathbf{v} = s\mathbf{n}$ where $\mathbf{n} \in S^{n-1}$ defines the directional heading.

We can therefore simply define $T$ in terms of a directional distribution, say $q$, for choosing some heading $\mathbf{n} \in S^{n-1}$. Specifically,

$$T(\mathbf{v},\mathbf{x},t) := \frac{q(\mathbf{n},\mathbf{x},t)}{s^{n-1}}\,, \tag{34}$$

where the $s^{n-1}$ factor results from moving between a distribution over $V$ to one over $S^{n-1}$. Subsequently, advection and diffusion tensors for (2) will be given by

$$\mathbf{a}(\mathbf{x},t) = s\mathbf{E}_q(\mathbf{x},t) = s\int_{S^{n-1}} \mathbf{n}q(\mathbf{n},\mathbf{x},t)d\mathbf{n}\,, \tag{35}$$

$$\mathbb{D}(\mathbf{x},t) = \frac{s^2}{\mu}\mathbb{V}_q(\mathbf{x},t) = \frac{s^2}{\mu}\int_{S^{n-1}} (\mathbf{n}-\mathbf{E}_q)(\mathbf{n}-\mathbf{E}_q)^T q d\mathbf{n}\,. \tag{36}$$

Notice that for the von-Mises and Fisher distributions discussed earlier, we have already computed expectation and variances: i.e. they are ready to be used.

**Example 2: (including external drift)** The above derivation can also be applied to the case of particles that are drifting in an external velocity field $\mathbf{b}(\mathbf{x},t) \in \mathbb{R}^n$, for example turtles transported in ocean currents or insects blown by the wind. If particles are inactive their heading is exactly the direction of the external flow field $\mathbf{b}(\mathbf{x},t)$, in which case the directional distribution used for the turning kernel would be a point measure

$$T(\mathbf{v},\mathbf{x},t) = \delta_{\mathbf{b}(\mathbf{x},t)}(\mathbf{v})\,.$$

Then, expectation and variances can be calculated as

$$\mathbf{E}_T(\mathbf{x},t) = \mathbf{b}(\mathbf{x},t) \quad \text{and} \quad \mathbb{V}_T(\mathbf{x},t) = 0\,.$$

The above macroscopic limit is a pure drift equation

$$u_t + \nabla\cdot(\mathbf{b}(\mathbf{x},t)u) = 0. \tag{37}$$

Note that the same equation arises if we simply assume that a force proportional to $\mathbf{b}$ acts on cells, where the cells have no inertia. In that case we also get a drift of the form $\mathbf{b}(x,t)$. For situations in which we have a population of actively navigating/moving particles immersed in an external velocity field we can simply combine the two cases of (35), (36) and (37) to obtain

$$u_t + \nabla\cdot((\mathbf{a}(\mathbf{x},t)+\mathbf{b}(\mathbf{x},t))u) = \nabla\nabla\colon\,(\mathbb{D}(\mathbf{x},t)u). \tag{38}$$

Indeed, this case was used to analyse sea turtle data in [43].

## 4 Applications to Cell/Animal Orientation Datasets

We illustrate the methodology through our two motivating applications. In each case we take as a starting point an individual-based description for oriented movement: an underlying velocity-jump process for the random walk. This initial description arises naturally, given our fundamental knowledge of particle behaviours: cells on fabricated substrates reveal alignment and orientation according to the substrate anisotropy

(Figure 1); datasets for turtles are based according to their mean swimming orientation when subjected to specific magnetic fields (Figure 2). We remark that in each application a two-dimensional approximation ($n = 2$) is reasonable: cells migrate across the two-dimensional substrate and the diving capabilities of young turtles restrict their movements to the ocean surface [10]. Simulation methods are provided in the Appendix.

The two applications differ not only in their field of study but also with respect to the "usable data". For cell movement we consider a tabulated summary of responses for distinct micro-ridge substrates, Table 1. This is data at a population-averaged level, and we do not have explicit data on each individual cell's orientating response. Nevertheless, we can still use this data to directly parametrise our model, which is done directly at the FAAD level that arises as a continuous approximation of the individual model. In the case of hatchling movements, a circular dataset is available for the mean heading of each tested turtle in samples exposed to distinct navigation fields. In this case, we can directly parametrise the von Mises distribution that describes an individual's orientation response, and subsequently scale to a macroscopic FAAD equation in order to collect population-level measurements.

## *4.1 Application A: Cell Movement on Microfabricated Structures*

The data of Jeon et al. [25] in Table 1 are at a population level: the mean $x$-velocity ($v_x \pm v_{x,error}$), mean $y$-velocity ($v_y \pm v_{y,error}$) and mean speed ($s \pm s_{error}$), where velocity components are measured according to absolute values. To relate these to the parametrisation of (2), we first remark on some particulars induced by the anisotropic arrangement. Firstly, the dominant drift velocity $\mathbf{a} = 0$, since the environment is essentially bidirectional and, on average, equal numbers of cells will be found travelling up or down (left or right). Secondly, the substratum is anisotropic but spatially homogeneous, and hence the diffusion tensor $\mathbb{D}$ is constant in space. Finally, anisotropies coincide with the coordinate axes, so $\mathbb{D}$ becomes a diagonal matrix

$$\mathbb{D} = \begin{pmatrix} \lambda_x & 0 \\ 0 & \lambda_y \end{pmatrix}, \tag{39}$$

with two eigenvalues $\lambda_x$ and $\lambda_y$.

Given that $\mathbb{D}$ is constant in space, the fully-anisotropic diffusion model becomes identical to the standard anisotropic diffusion equation:

$$u_t = \nabla \cdot \mathbb{D} \nabla u\,. \tag{40}$$

Hence we can exploit results relating to the above. Firstly, the fundamental solution of (40) is the Gaussian distribution with covariance matrix $\mathbb{D}$:

$$u(\mathbf{x},t) = \frac{1}{2\pi t\sqrt{\mathrm{Det}\mathbb{D}}} \exp\left(-\frac{1}{4t}\mathbf{x}^T \mathbb{D}^{-1}\mathbf{x}\right) \tag{41}$$

(in two spatial dimensions), where the set

$$E_c := \{\mathbf{x} : \mathbf{x}^T \mathbb{D}^{-1}\mathbf{x} = c\}$$

gives the set of locations for which there is an equal probability of finding a random mover that started at the origin. This set defines a *diffusion ellipse*, with semi-axes of lengths $\sqrt{\lambda_x}$ and $\sqrt{\lambda_y}$ respectively, and provides one way to graphically visualise the anisotropy of $\mathbb{D}$. A second method is the *diffusion peanut*, which is the

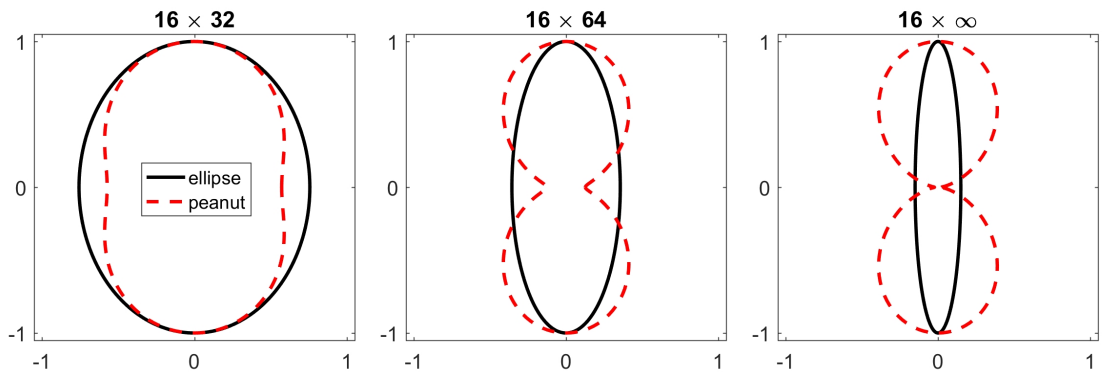


**Fig. 6** Diffusion ellipses (black solid line) and peanuts (red-dashed line) representing the anisotropic cell migration for the $16\times 32$, $16\times 64$ and $16\times\infty$ micro-ridge arrangements, see Table 2. Note that we renormalise the longer axes to aid comparison between their respective shapes.

image of the map $\mathbf{w}\mapsto \mathbf{w}^T\mathbb{D}\mathbf{w}$ for $\mathbf{w}\in S^1$, and relates to the mean-squared displacement in direction $\mathbf{w}$, $\sigma^2_{\mathbf{w}}$, via $\sigma^2_{\mathbf{w}} = 2t\mathbf{w}^T\mathbb{D}\mathbf{w}$ [45]. This gives rise to the *apparent diffusion coefficient in direction* $\mathbf{w}$,

$$\mathrm{ADC}_{\mathbf{w}} := \frac{\sigma^2_{\mathbf{w}}}{2t} = \mathbf{w}^T\mathbb{D}\mathbf{w}\,.$$

In particular, given coordinate directions $(1,0)^T$ and $(0,1)^T$, we find that the mean squared displacements in $x-$ and $y-$directions will be $2t\lambda_x$ and $2t\lambda_y$ respectively. This provides the key for using the data in Table 1: given the mean velocities in $x$ and $y$ directions and taking a unit time step of 1 minute, we convert to mean displacements for the $x$ and $y$ directions and in turn estimate the $\lambda$'s in (39), the values of which are listed in Table 2 for each experimental setting. To illustrate some of the anisotropies graphically, we plot diffusion ellipses and peanuts for the three cases $16\times 32$, $16\times 64$ and $16\times\infty$ in Figure 6. As the structure is stretched along the $y-$ direction we observe progressively thinned-out ellipses/pinched peanuts, reflecting restricted movement along this axis.

For turning rates of the order of $2.5/\mathrm{min}$ and a tracking timeframe of 400 minutes, each cell turns on average 1000 times across its track. Given an average speed of 0.5 $\mu$m/min, each particle travels about 200 $\mu$m in this timeframe, suggesting this to be a suitably macroscopic scale. We subsequently plot solutions to the FAAD model on this spatial and temporal scale, plotting the evolving distribution for 10 individuals presumed to have started at the origin. Exploiting the spatially uniform nature of the environment, solutions will simply be governed by the fundamental solution (41), which we plot in Figure 7 at $t = 100$ and $t = 400$ for the same three cases $16\times 32$, $16\times 64$ and $16\times\infty$. Consistent with the diffusion ellipses, the highest degree of environmental anisotropy generates a quasi-one dimensional spread of the cells along the $y-$axis. We note that there is no direct information in [25] that allows us to directly compare these plots to their data, and therefore this represents a prediction of the expected population distribution.

We can turn the argument full circle and use the measured data to estimate cell movement parameters that would be required in the underlying velocity-jump process: speed $s$, turning rate $\mu$, and concentration parameter $\kappa$ of the bimodal von-Mises distribution (14). We should note that this is predicated on an *assumption* of the individual-level behaviour: i.e. that cells orient according to a bimodal von-Mises distribution. In the

| Case ($\mu$m$\times\mu$m) | Ridge height ($\mu$m) | Speed $\pm$ error ($\mu$m/min) | $\lambda_x\pm$ error ($\mu$m$^2$/min) | $\lambda_y\pm$ error ($\mu$m$^2$/min) | Turning rate (/min) | Anisotropy Parameter |
|---|---|---|---|---|---|---|
| 12 x 24 | 3 | 0.78±0.027 | 0.072±0.0057 | 0.17±0.015 | 2.53 | 2.57 |
| 12 x 48 | 3 | 1.01±0.045 | 0.039±0.0039 | 0.41±0.041 | 2.29 | 10.79 |
| 12 x ∞ | 3 | 0.59±0.029 | 0.0032±0.00040 | 0.16±0.016 | 2.17 | 49.49 |
| 16 x 32 | 3 | 0.9±0.03 | 0.12±0.010 | 0.21±0.017 | 2.48 | 1.96 |
| 16 x 64 | 3 | 1.0±0.039 | 0.048±0.0047 | 0.38±0.033 | 2.34 | 8.32 |
| 16 x ∞ | 3 | 0.84±0.0072 | 0.0072±0.00080 | 0.32±0.029 | 2.15 | 44.84 |
| 24 x 48 | 3 | 0.55±0.027 | 0.034±0.0039 | 0.088±0.010 | 2.47 | 2.89 |
| 24 x 96 | 3 | 0.58±0.022 | 0.020±0.0024 | 0.12±0.0098 | 2.40 | 6.42 |
| 24 x ∞ | 3 | 0.52±0.028 | 0.0072±0.00084 | 0.12±0.013 | 2.20 | 16.47 |
| 12 x 24 | 10 | 0.65±0.026 | 0.055±0.0053 | 0.11±0.011 | 2.63 | 2.10 |
| 12 x 48 | 10 | 0.83±0.046 | 0.016±0.0023 | 0.29±0.033 | 2.25 | 18.28 |
| 12 x ∞ | 10 | 0.61±0.032 | 0.00081±0.00012 | 0.18±0.019 | 2.05 | 224.22 |
| control | 0 | 0.63±0.025 | 0.072±0.0072 | 0.085±0.014 | 2.53 | 0.83 |

**Table 2** Speed and diffusion coefficients $\lambda_x$ and $\lambda_y$ from the data from Jeon et al. [25]. We also list the values for the turning rate $\mu$, and the concentration parameter $\kappa$ of a corresponding bi-modal von-Mises distribution.

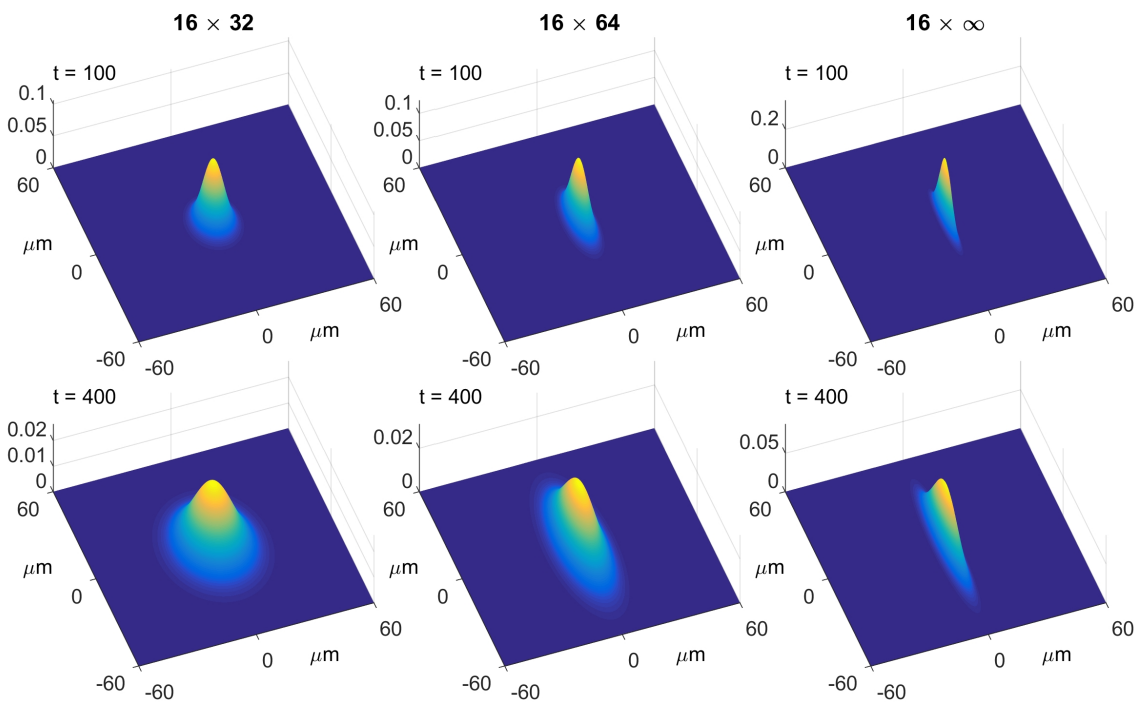


**Fig. 7** Population distributions $u(\mathbf{x},t)$ plotted at (top row) $t = 100$ and (bottom row) $t = 400$ for 10 cells initiated at $\mathbf{x} = \mathbf{0}$.

absence of specific individual-level data, this is of course impossible to state with certainty, yet it is never-

theless instructive to show how we can "reverse the process".

Recall that, given the symmetric/bidirectional scenario, the drift velocity $\mathbf{a} = 0$ and the macroscopic model becomes the pure fully anisotropic diffusion equation

$$u_t = \nabla\nabla : (\mathbb{D}u),$$

with diffusion tensor from (36)

$$\mathbb{D} = \frac{s^2}{\mu}\mathbb{V}_q = \frac{s^2}{2\mu}\left(1 - \frac{I_2(\kappa)}{I_0(\kappa)}\right)\mathbb{I}_2 + \frac{s^2}{\mu}\frac{I_2(\kappa)}{I_0(\kappa)}\boldsymbol{\nu}\boldsymbol{\nu}^T. \tag{42}$$

For now let us write the diffusion tensor in (42) as

$$\mathbb{D} = k_1\mathbb{I}_2 + k_2\boldsymbol{\nu}\boldsymbol{\nu}^T, \qquad k_1 = \frac{s^2}{2\mu}\left(1 - \frac{I_2(\kappa)}{I_0(\kappa)}\right), \qquad k_2 = \frac{s^2}{\mu}\frac{I_2(\kappa)}{I_0(\kappa)}. \tag{43}$$

Since the primary direction of anisotropy is in the $y$-direction, we have $\boldsymbol{\nu} = (0,1)^T$ and can explicitly compute

$$\mathbb{D} = \begin{pmatrix} k_1 & 0 \\ 0 & k_1 + k_2 \end{pmatrix} = \begin{pmatrix} \lambda_x & 0 \\ 0 & \lambda_y \end{pmatrix},$$

where we employed (39) for the second equality. Therefore, we obtain two equations relating $k_1, k_2$ and $\lambda_x, \lambda_y$:

$$k_1 = \lambda_x \qquad k_1 + k_2 = \lambda_y.$$

Using the expressions for $k_1$ and $k_2$ in (43) we find $\mathrm{tr}\mathbb{D} = \lambda_x + \lambda_y = \frac{s^2}{\mu}$, which gives

$$\mu = \frac{s^2}{\lambda_x + \lambda_y}. \tag{44}$$

The corresponding values for the turning rate $\mu$ are listed in Table (2). Furthermore we can use the previous relations to compute

$$\frac{I_2(\kappa)}{I_0(\kappa)} = \frac{\mu(\lambda_y - \lambda_x)}{s^2}. \tag{45}$$

Determining concentration (or anisotropy) parameter $\kappa$ demands inverting the ratio of modified Bessel functions $I_2(\kappa)/I_0(\kappa)$, a monotonically increasing function from 0 to 1 for $\kappa \in [0,\infty)$. We use Wolfram Alpha to invert this function for our data and list the corresponding values in Table (2).

The turning rate $\mu$ is surprisingly consistent between the different experiments, which may reflect that this parameter is (relatively) independent of the form of the substratum (for example, determined mainly by intracellular factors). The anisotropy parameter $\kappa$, however, varies over several orders of magnitude with the most anisotropic cases corresponding to those without ridges in the $x$-direction, as expected. Graphical illustrations of the bimodal von Mises distribution for the three cases $16 \times 32$, $16 \times 64$ and $16 \times \infty$ are provided in Figure 8. Higher ridges (10 $\mu$m) offer even more guidance and, consequently, larger anisotropy: including an extreme of $\kappa = 224$. This upper value effectively reduces the bimodal von Mises distribution to a pair of Delta functions in opposite directions, so that movement is almost completely confined to the one-dimensional $y-$direction.

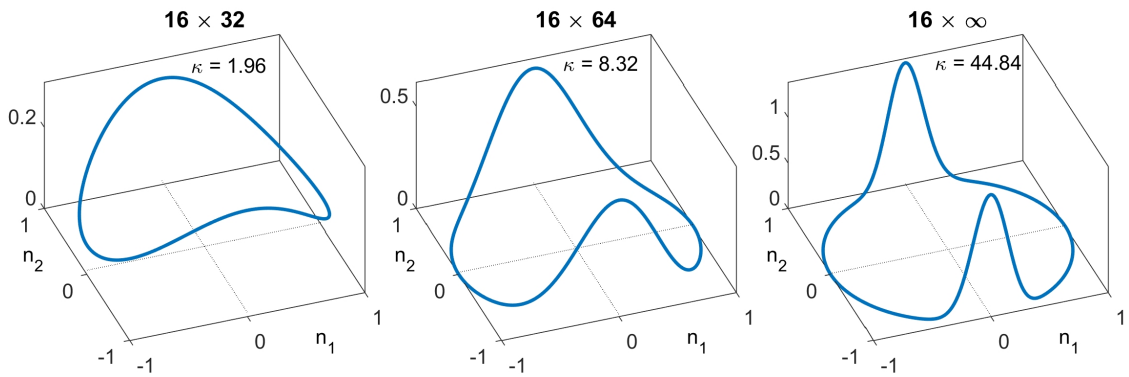


**Fig. 8** Bimodal von Mises distributions for the turning distributions of stochastic velocity-jump random walks corresponding to the macroscopic cases in Figure 7.

## *4.2 Application B: Magnetic Navigation in Loggerhead Hatchlings*

Our second application considers hatchling loggerhead turtle navigation, investigating the extent to which oriented swimming keeps them within the relative safety of the North Atlantic Gyre. Specifically, we extend the agent-based simulation study of [51], exploiting the computational advantages of the FAAD model to investigate how different amounts of oriented swimming help to maintain turtle trajectories. We specifically focus on two critical regions of the Gyre as follows.

- (NE) a north east Gyre location corresponding to a "corridor" along its northeastern sector, the region where it breaks into northerly (perilous) and southerly moving streams. We center this region on the point marked 3 in Figure 2, with its corresponding dataset providing the parameters for orientation.
- (SW) a south west Gyre location corresponding to a region of the Carribean, where the Gyre branches into a more northerly stream that remains within the Gyre, or continues west into the Gulf of Mexico. We center this region on point 7 in Figure 2, with its corresponding dataset providing the parameters for orientation.

In each case we quantitatively assess the extent to which hatchling turtles that are continuously immersed at some point inside (NE) or (SW) tend to maintain a trajectory within the Gyre. Specifically, for each region (NE) and (SW) we numerically solve the FAAD equation, as extended to incorporate both an additional drift (as derived above, see equation (38)) due to currents and a constant (in time) source representing hatchlings entering the region under investigation. Specifically, defining $u(\mathbf{x},t)$ to be the hatchling turtle density, we solve

$$u(\mathbf{x},t)_t + \nabla\cdot((\mathbf{a}(\mathbf{x},t)+\mathbf{b}(\mathbf{x},t))u(\mathbf{x},t)) = \nabla\nabla : (\mathbb{D}(\mathbf{x},t)u(\mathbf{x},t)) + \gamma\delta_{\mathbf{x}_0}(\mathbf{x}), \tag{46}$$

where, in addition to previous definitions, $\gamma$ represents the rate at which new hatchlings enter the system and $\delta_{\mathbf{x}_0}$ is the 2D Dirac delta function. The point $\mathbf{x}_0$ defines the "immersion site" and we set $\mathbf{x}_0 = (25^\circ W, 44.5^\circ N)$ for (NE) and $\mathbf{x}_0 = (56.5^\circ W, 8^\circ N)$ for (SW), respectively denoting points upwards of the general current direction for the regions. Encountered currents $\mathbf{b}(\mathbf{x},t)$ can vary considerably over time, and we therefore inject hatchlings continuously into the corridor across a full calender year (taken to be 2016). Our restriction to the two-dimensional ocean surface follows from the poor diving abilities of young marine turtles: a maximum

dive of the order of 1-2 metres for loggerhead hatchlings [10].

We define a "success" and a "failure" boundary for each region, removing turtles if they hit either of these boundaries and tracking over time the total numbers that have done so. In the context of the continuous model, this corresponds to setting absorbing boundary conditions along two boundaries. For the (NE) region we define the success boundary along the 42.5°N line and the failure boundary along 46.5°N line; the more northerly line represents turtles moving towards cooler waters and straying from the southerly shifting Gyre. For (SW) the success boundary is set along 18°N line and the failure boundary marked by 64.5°W; success is implied by a northerly shift with the Gyre, while failure is marked by a westward shift towards the Gulf of Mexico. Of course, the lack of any data makes any such notion of success or failure moot and we cannot equate these boundaries with survival probabilities: they simply provide a proxy to track the tendency to remain within the Gyre.

To close the computational regions we consider two further boundaries with reflective boundary conditions associated with them, so that there is no net loss across these boundaries. For (NE) we consider the lines 28°W/12°W, and for (SW) the lines 54.5°W/8°N. Note that these lines are all reasonably far from the initial injection site such that, in practice, the vast majority of turtles end up becoming absorbed by one of the success/failure boundaries before hitting one of the reflective boundaries.

### 4.2.1 Data and parametrisation

The model demands two specific components that can be drawn from biological data: the ocean currents $\mathbf{b}(\mathbf{x},t)$ for the passive drift vector field and navigation/movement parameters for hatchling active movement. Velocity fields for ocean currents are obtained from HYCOM (the global HYbrid Coordinate Ocean Model, [6]), an ocean forecasting model forced by wind speed, heat flux and numerous other factors that has been subsequently assimilated with field measurements (from satellites, floats, moored buoys etc) to generate post-validated output. The resolution of HYCOM data ($1/12^\circ$ and day to day) allows it to reproduce both the large scale persistent currents and localised phenomena such as eddies. Note that the surface/near-surface swimming behaviour of young tutles allows us to restrict to the (2D) uppermost layer of HYCOM datasets. HYCOM data for each of regions (NE) and (SW) was downloaded from `http://pdrc.soest.hawaii.edu/data/data.php`, accessed during June/July 2017. Note that for computations, HYCOM data has been interpolated from its native resolutions ($1/12^\circ$ and day-day) to the spatial/temporal resolution required by the numerical code via standard linear interpolation schemes.

Defining the active movement component to motion requires specifying the speed/turning rate $(s,\lambda)$ parameters and the concentration/dominant direction $(\kappa,\boldsymbol{\nu})$ parameters demanded by the von Mises distribution. Hatchlings are capable of sustaining speeds of 0.72 km/hr (see [51] and references therein) and, based on this, we suppose the average daily swim length varies from 0-10 km/day, corresponding to between 0 and $\sim$14 hours per day of active swimming. Of course, whether a hatchling would be capable of maintaining active swimming at the upper end of this spectrum is somewhat debatable. For the turning rate, we assume a value of 50 per day, although it is noted that modifying this parameter has very little bearing on the overall results. Given this turning rate and assuming each turtle remains in the simulated region for the order of 100 days, we obtain an average of 5000 turns per trajectory. For average swimming speeds ranging between 0-10 km/day, turtles swim up to 1000 km over the simulation timecourse, implying spatial scales of the order 100-1000 km as suitably macroscopic. We remark that the comparisons between the individual and continuous

simulations suggest the veracity of the continuous limit as a suitable approximation.

Concentration parameters/dominant directions can be drawn directly from the hatchling orientation datasets illustrated in Figure 2. For region (NE) we utilise the dataset indicated by position 3: fitting a von Mises distribution via standard methods (e.g. see [2]) allow us to obtain estimates $\kappa_{NE} \approx 0.874$ and $\boldsymbol{\nu}_{NE} \approx (0.307, -0.952)$, the latter representing a true bearing of $162^\circ$. The region (SW) employs position 7 and yields $\kappa_{SW} \approx 0.797$ and $\boldsymbol{\nu}_{SW} \approx (0.070, 0.998)$, representing a true bearing of $4^\circ$. We assume these values are constant in space and time over the respective regions.

### 4.2.2 Results

In Figure 9 we compare the density distribution predicted by the parametrised FAAD model (46) with a particle distribution obtained through individual-based simulations of the stochastic velocity-jump process. The close correlation between the continuous density distribution (as reflected by the colormap) and the distribution of individual particles (white dots) indicates that the FAAD model provides a highly acceptable approximation for the turtle distribution. Further simulations (not shown) confirm this close correspondence, and we therefore exploit the FAAD model for its computational advantages in the subsequent simulations.

Figure 10 compares density distributions for the same region at the same time points under three choices for the amount of active swimming: 0 km/day (i.e. only passive drifting occurs), 2 km/day and 10 km/day. A shift towards a greater amount of active swimming has a clear impact on the density distribution, pushing it in an expected southerly direction such that a greater density becomes absorbed by the "success" boundary.

Finally, we plot the results from a more extended analysis, following a parameter sweep for each of the two regions, classifying the data obtained in terms of the following simple "success measure":

$$\text{Success at time T} = \frac{\text{Total density hitting success boundary by time T}}{\text{Total density hitting success and failure boundaries by time T}}.$$

The above clearly approaches 1 for a successful population and 0 for an unsuccessful population. In the simulations here we set $T = 500$ for a population continuously released at $\mathbf{x}_0$ from $t = 0$ (midnight, 01/01/2016) to the end of 2016 ($t = 366$); the continuation until $T = 500$ ensures that by the end of the simulation only a negligible fraction of the released population has failed to hit one of the absorbing boundaries. Simulations are plotted in Figure 11 for each of the two regions, under a range of daily active swimming distances and for three values of the concentration parameter: the value obtained by the data fitting and perturbations of $\times 2$ and $\times 1/2$ these values. The simulations clearly show that increasing the amount of active swimming, or increasing the certainty of orientation, nudges a greater proportion of the population towards the successful boundary, supporting the hypothesis that oriented responses can help maintain hatchling movement within the Gyre (e.g. [28, 51]). Extensions of the study to consider movement throughout the full circulatory path would allow more detailed evaluations into the extent to which oriented swims aid route maintenance: we remark that this would be a focus for a future study and refer to [28] for such an analysis for an individual-based model.

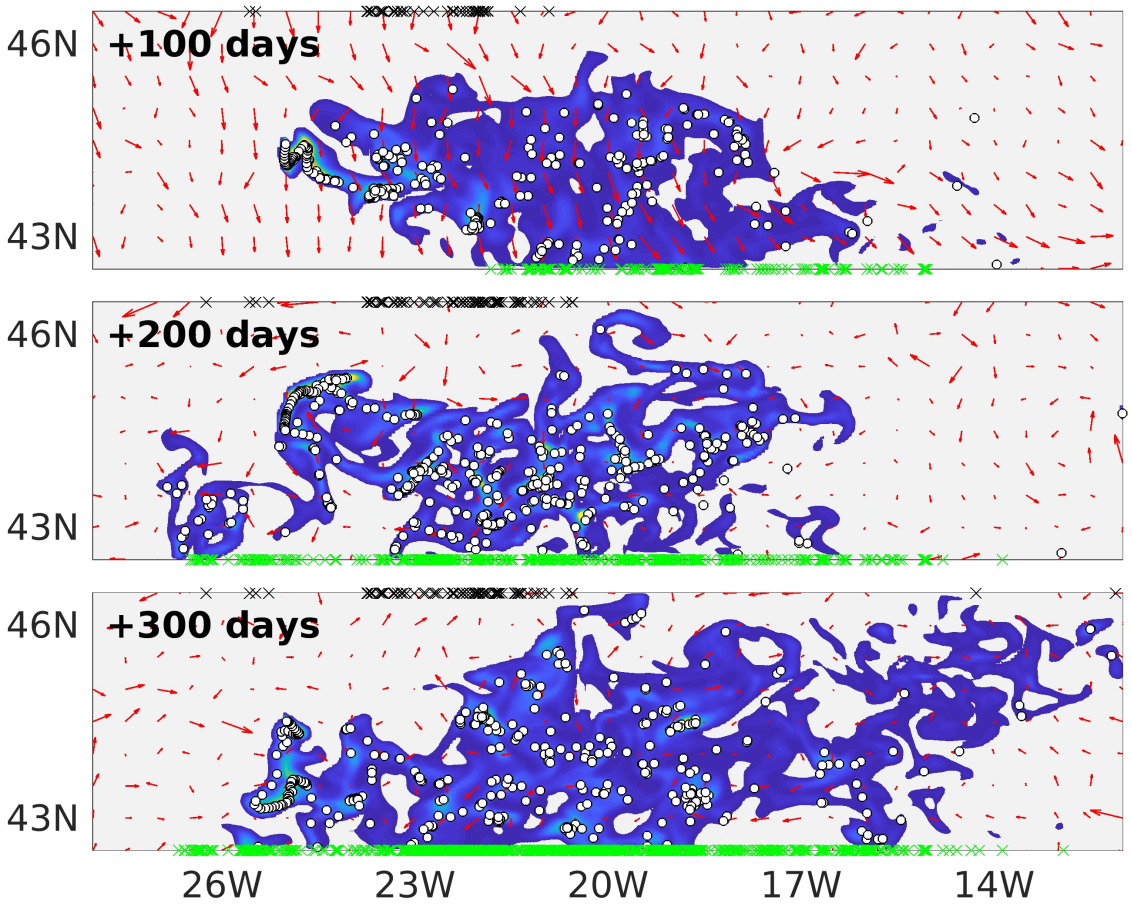


**Fig. 9** Comparison between the FAAD model (46) and individual-based stochastic simulations of the velocity-jump model for the problem of North Atlantic turtle hatchling movement. In each frame we plot both the continuous population density distribution $u(\mathbf{x},t)$ (reflected by the colour map, where grey indicates negligible density and blue to yellow reflects increasing density) and the individual dots generated by the velocity-jump simulations. Here, top and bottom boundaries respectively define the "failure" and "success" boundaries, and the individual particles are colour coded according to whether they are still moving (white dots) or have hit either the failure (black crosses) or success (green crosses) boundary. Underlying ocean currents are indicated by the red arrows. For this simulation we use region (NE) and release particles continuously from position $\mathbf{x}_0 = (25^\circ W, 44.5^\circ N)$ with $\gamma = 5$/day. The total daily swim is set at $s = 2$ km/day, with $\lambda = 50$/day, $\kappa_{NE} \approx 0.874$ and $\boldsymbol{\nu}_{NE} \approx (0.307, -0.952)$. Note that the von Mises distribution for these values is visualised by the dashed red line in the inset figure to the left hand frame of Figure 11. Simulations (in terms of ocean currents utilised) start on 01/01/2016 (midnight) with solutions displayed on the days following as indicated.

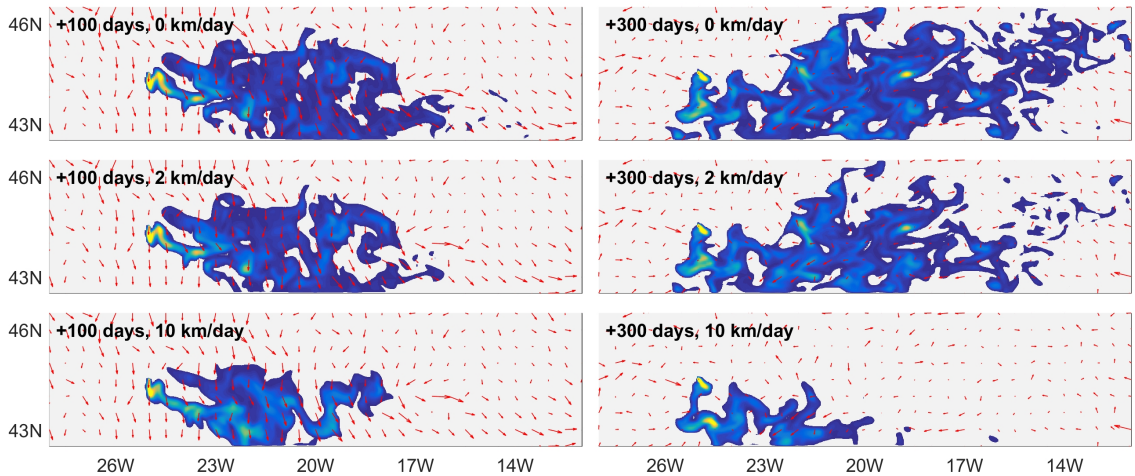


**Fig. 10** Comparison of population density distributions under varying amounts of active swimming per day. In each frame we plot the turtle density distribution (color density map, as described in Figure 9) at the two separate times (left) $+100$ days and (right) $+300$ days for (top row) $s = 0$ km/day, (middle row) $s = 2$ km/day and (bottom row) $s = 10$ km/day. The strength and direction of ocean currents is indicated by the red arrows. All other parameters and details as in Figure 9.

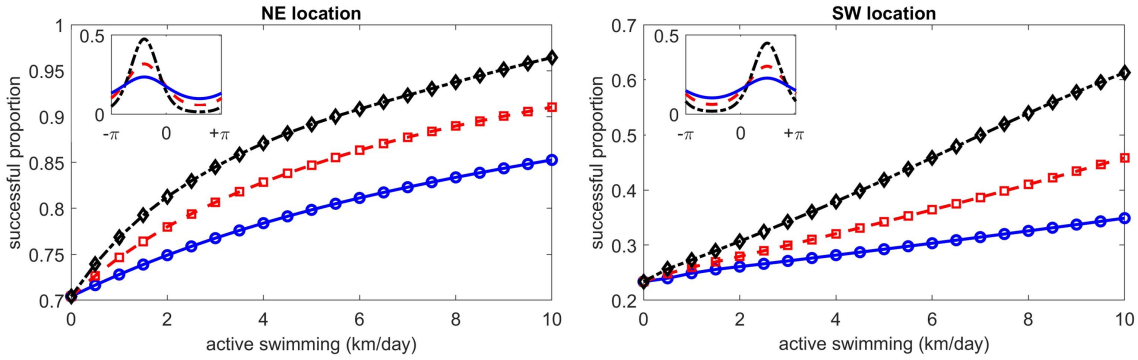


**Fig. 11** Success is plotted as a function of daily swimming distance for the two regions and for different concentration parameters. All other parameters and details as in Figure 9. Red dashed line indicates a choice of $\kappa$ as taken directly from the data fitting, with blue solid and black dot-dashed respectively showing choices of $\times 2$ and $\times 1/2$ these values. Insets plot the corresponding von Mises distributions used for each simulation set.

## 5 Conclusions

In this chapter we have described the use of fully-anisotropic advection-diffusion models as a way of modelling animal and cell movement behaviour. We have described the derivation of these models from two fundamental stochastic random walks, position-jump and velocity-jump processes, thereby connecting the macroscopic parameters and terms to the statistical inputs at the individual level. Utilising two distinct datasets, we have shown how the models can be parametrised either directly at the population level, or

by starting at the individual/stochastic random walk model. Beyond the applications presented here, we note that similar methods have been applied in a number of other applications in ecology and cell movement, including seismic-line following behaviour of wolves and caribou populations [33, 22], butterfly hilltopping [44], and anisotropic glioma growth [45, 56].

**Acknowledgements**: KJP would like to acknowledge the Politecnico di Torino for the Visiting Professor award. TH is grateful to the Politecnico di Torino for their hospitality. TH is supported through a discovery grant of the Natural Science and Engineering Research Council of Canada (NSERC), an Oberwolfach (MFO) Simons Foundation Visiting Professorship and the Edinburgh Mathematical Society.

## References

1. Alt, W.: Biased random walk model for chemotaxis and related diffusion approximation. J. Math. Biol. **9**, 147–177 (1980)
2. Batschelet, E.: Circular Statistics in Biology. Academic Press, London (1981)
3. Bellomo, N., Schiavo, M.: Lecture Notes on the Mathematical Theory of Generalized Boltzmann Methods. World Scientific, Singapore (2000)
4. Berens, P.: Circstat: a MATLAB toolbox for circular statistics. J. Stat. Softw. **31**, 1–21 (2009)
5. Berg, H.: Random Walks in Biology. Princeton University Press (1983)
6. Bleck, R.: An oceanic general circulation model framed in hybrid isopycnic-cartesian coordinates. Ocean Mod. **4**, 55–88 (2002)
7. Cagnacci, F., Boitani, L., Powell, R.A., Boyce, M.S.: Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges. Phil. Trans. R. Soc. B **365**, 21572162 (2010)
8. Cercignani, C., Illner, R., Pulvirenti, M.: The Mathematical Theory of Diluted Gases. Springer, New York (1994)
9. Codling, E.A., Plank, M.J., Benhamou, S.: Random walk models in biology. J. Roy. Soc. Interface **5**, 813–834 (2008)
10. Davenport, J. and Clough, W.: Swimming and diving in young loggerhead sea turtles (*Caretta caretta L.*). Copeia, **1986**, 53–57 (1986)
11. Dawes, A., Iron, D.: Cortical geometry may influence placement of interface between par protein domains in early caenorhabditis elegans embryos. J. Theor. Bio. **333**, 27–37 (2013)
12. Deutsch, A., Dormann, S.: Cellular Automaton Modeling of Biological Pattern Formation: Characterization, Applications, and Analysis. Birkaeuser, Boston (2005)
13. Dickinson, R.B., Guido, S., Tranquillo, R.T.: Biased cell migration of fibroblasts exhibiting contact guidance in oriented collagen gels. Ann. Biomed. Eng. **22**, 342–356 (1994)
14. Dunn, G.A., Heath, J.P.: A new hypothesis of contact guidance in tissue cells. Exp. Cell Res. **101**, 1–14 (1976)
15. Fuxjager, M.J., Eastwood, B.S., Lohmann, K.J.: Orientation of hatchling loggerhead sea turtles to regional magnetic fields along a transoceanic migratory pathway. J. Exp. Biol. **214**, 2504–2508 (2011)
16. Gritsenko, P., Ilina, O., Friedl, P.: Interstitial guidance of cancer invasion. J. Pathol. **226**, 185–199 (2012)
17. Gritsenko, P., Leenders, W., Friedl, P.: Recapitulating in vivo-like plasticity of glioma cell invasion along blood vessels and in astrocyte-rich stroma. Histochem. Cell Biol. (2017). doi: 10.1007/s00418-017-1604-2
18. Hadeler, K., Hillen, T., Lutscher, F.: The Langevin or Klein-Kramers approach to biological modeling. Math. Models Meth. Appl. Sci. **14**(10), 1561–1583 (2004)
19. Hanahan, D., Weinberg, R.: Hallmarks of cancer: The next generation. Cell **144**, 646–674 (2011)
20. Hillen, T.: $M^5$ mesoscopic and macroscopic models for mesenchymal motion. J. Math. Biol. **53**, 585–616 (2006)
21. Hillen, T., Othmer, H.: The diffusion limit of transport equations derived from velocity jump processes. SIAM J. Appl. Math. **61**, 751–775 (2000)
22. Hillen, T., Painter, K.J.: Transport models for movement in oriented habitats and anisotropic diffusion. In: Lewis,M., Maini,P., Petrovskii,S. (Eds.), Dispersal, Individual Movement and Spatial Ecology: A Mathematical Perspective. Springer,Heidelberg. p. 46 (2013)
23. Hillen, T., Painter, K.J., Swan, A.C., Murtha, A.D.: Moments of von Mises and Fisher distributions and applications. Math. Biosci. & Eng **14**, 673–694 (2017)
24. Hundsdorfer, W., Verwer, J.G.: Numerical solution of time-dependent advection-diffusion-reaction equations, vol. 33. Springer Science & Business Media (2003)

25. Jeon, H., Hidai, H., Hwang, D.J., Healy, K.E., Grigoropoulos, C.P.: The effect of micronscale anisotropic cross patterns on fibroblast migration. Biomaterials **31**, 4286–4295 (2010)
26. Keener, J., Sneyd, J.: Mathematical Physiology. Springer (1994)
27. Lohmann, K.J., Cain, S.D., Dodge, S.A., Lohmann, C.M.F.: Regional magnetic fields as navigational markers for sea turtles. Science **294**, 364–366 (2001)
28. Lohmann, K.J., Putman, N.F., Lohmann, C.M.F.: The magnetic map of hatchling loggerhead sea turtles. Curr. Opin. Neurobiol. **22**, 336–342 (2012)
29. Luschi, P.: Long-distance animal migrations in the oceanic environment: orientation and navigation correlates. ISRN Zool. (2013)
30. Lutscher, F., Pachepsky, E., Lewis, M.: The effect of dispersal patterns on stream populations. SIAM J. Appl. Math. **65**, 1305–1327 (2005)
31. Mardia, K., Jupp, P.: Directional Statistics. Wiley and Sons (2000)
32. McKenzie, H., Lewis, M., Merrill, E.: First passage time analysis of animal movement and insights into the functional response. Bull. Math. Biol. **71**, 107–129 (2009)
33. McKenzie, H.W., Merrill, E.H., Spiteri, R.J., Lewis, M.A.: How linear features alter predator movement and the functional response. Interface focus **2**, 205–216 (2012)
34. Moorcroft, P., Lewis, M.: Mechanistic Home Range Analysis. Princeton University Press, Princeton (2006)
35. Murray, J.: Mathematical Biology. I: An Introduction, 3rd edn. Springer-Verlag, New York (2002)
36. Murray, J.D.: Mathematical biology II: Spatial models and biochemical applications, Springer-Verlag, New York (2003)
37. Okubo, A., Levin, S.: Diffusion and Ecological Problems: Modern Perspectives. Springer (2002)
38. Othmer, H., Dunbar, S., Alt, W.: Models of dispersal in biological systems. J. Math. Biol. **26**, 263–298 (1988)
39. Othmer, H.G., Stevens, A: Aggregation, blowup, and collapse: the ABC's of taxis in reinforced random walks. SIAM J. Appl. Math., **57**, 1044-1081 (1997).
40. Othmer, H., Hillen, T.: The diffusion limit of transport equations II: Chemotaxis equations. SIAM J. Appl. Math. **62**, 1122–1250 (2002)
41. Othmer, H.G., Xue, C.: The mathematical analysis of biological aggregation and dispersal: progress, problems and perspectives. In: Lewis,M., Maini,P., Petrovskii,S. (Eds.), Dispersal, Individual Movement and Spatial Ecology: A Mathematical Perspective. Springer,Heidelberg. 79–127 (2013)
42. Painter, K.J.: Modelling migration strategies in the extracellular matrix. J. Math. Biol. **58**, 511–543 (2009)
43. Painter, K.J., Hillen, T.: Navigating the flow: Individual and continuum models for homing in flowing environments. Royal Soc. Interface **12**, 20150,647 (2015)
44. Painter, K.J.: Multiscale models for movement in oriented environments and their application to hilltopping in butterflies. Theor. Ecol. **7**, 53–75 (2014)
45. Painter, K.J., Hillen, T.: Mathematical modelling of glioma growth: the use of diffusion tensor imaging (DTI) data to predict the anisotropic pathways of cancer invasion. J. Theor. Biol. **323**, 25–39 (2013)
46. Patlak, C.: Random walk with persistence and external bias. Bull. Math. Biophys. **15**, 311–338 (1953)
47. Perthame, B.: Transport Equations in Biology. Birkhäuser (2007)
48. Preziosi, L. (ed.): Cancer Modelling and Simulation. Chapman Hall/CRC Press (2003)
49. Provenzano, P.P., Eliceiri, K.W., Campbell, J.M., Inman, D.R., White, J.G., Keely, P.J.: Collagen reorganization at the tumor-stromal interface facilitates local invasion. BMC medicine **4**, 38 (2006)
50. Putman, N.F., Endres, C.S., Lohmann, C.M.F., Lohmann, K.J.: Longitude perception and bicoordinate magnetic maps in sea turtles. Curr. Biol. **21**, 463–466 (2011)
51. Putman, N.F., Verley, P., Shay, T.J., Lohmann, K.J.: Simulating transoceanic migrations of young loggerhead sea turtles: merging magnetic navigation behavior with an ocean circulation model. J. Exp. Biol. **215**, 1863–1870 (2012)
52. Saxton, M.J., Jacobson, K.: Single-particle tracking: applications to membrane dynamics. Ann. Rev. Biophys. & Biomol. Struct. **26**, 373–399 (1997)
53. Sobel, D.: Longitude: The true story of a lone genius who solved the greatest scientific problem of his time. Bloomsbury Publishing USA (1995)
54. Stevens, A.: The derivation of chemotaxis-equations as limit dynamics of moderately interacting stochastic many particle systems. SIAM J. Appl. Math. **61**(1), 183–212 (2000)
55. Stevens, A., Othmer, H.G.: Aggregation, blowup, and collapse: the ABC's of taxis in reinforced random walks. SIAM J. Appl. Math. **57**, 1044–1081 (1997)
56. Swan, A., Hillen, T., Bowman, J.C., Murtha, A.D.: A patient-specific anisotropic diffusion model for brain tumour spread. Bull. Math. Biol. pp. 1–33 (2017)
57. Turchin, P.: Quantitative Analysis of Movement. Sinauer Assoc., Sunderland (1998)
58. Weickert, J.: Anisotropic diffusion in image processing. Teubner, Stuttgart (1998)

59. Wolf, K., Müller, R., Borgmann, S., Bröcker, E.B., Friedl, P.: Amoeboid shape change and contact guidance: T-lymphocyte crawling through fibrillar collagen is independent of matrix remodeling by MMPs and other proteases. Blood **102**, 3262–3269 (2003)

# Appendix: Numerical methods

## *Stochastic Velocity-Jump Process*

The stochastic random walk simulations assume each individual performs a velocity-jump random walk in either a static (cell movement) or flowing (turtles) medium. Particle motion therefore derives from an oriented and active movement component that describes the individual's self motility (crawling, swimming, flying etc), the details of which are encoded in the velocity-jump random walk, and a passive drift due to movement of the medium (e.g. air or water flow). The passive drift is described by a velocity vector field $\mathbf{b}(\mathbf{x},t)$ ($\mathbf{x}$ is position and $t$ is time) that could be either imposed (e.g. obtained from public-domain datasets) or separately modelled (e.g. Navier-Stokes equation). Note that we implicitly assume that the individuals have negligible impact on the flow of the surrounding medium.

For an individual $i$ at position $\mathbf{x}_i(t)$ and time $t$, travelling with active velocity $\mathbf{v}_i(t) = s(\cos\alpha_i(t), \sin\alpha_i(t))$ where angle $\alpha_i(t)$ denotes the active heading, then at time $t+\Delta t$ (where $\Delta t$ is small) we have:

$$\begin{aligned} \mathbf{x}_i(t+\Delta t) &= \mathbf{x}_i(t) + \Delta t(\mathbf{v}_i(t) + \mathbf{b}(t,\mathbf{x}_i))\,; \\ \mathbf{v}_i(t+\Delta t) &= \begin{cases} \mathbf{v}_i'(t+\Delta t) & \text{with probability } \lambda\Delta t\,, \\ \mathbf{v}_i(t) & \text{otherwise}\,. \end{cases} \end{aligned} \tag{47}$$

where $\mathbf{v}_i'(t+\Delta t)$ is the new velocity chosen at time $t+\Delta t$ if a reorientation has occurred, randomly chosen according to the given probability distribution for the turning kernel of the velocity jump random walk.

The time discretisation $\Delta t$ used in simulation is suitably small, in the sense that simulations conducted with smaller timesteps generate near identical results. For the selection of new active headings via the von Mises distribution we employ code (circ_vmrnd.m) from the circular statistics toolbox [4]. Currents and the inputs required for the active heading choice are interpolated from the native spatial/temporal resolutions in the saved variables to the individual particle's continuous position $\mathbf{x}$ and time $t$ via a simple linear interpolation scheme.

## *Continuous Model*

As described earlier, moment closure analysis for the velocity-jump random walk generates a continuous model of FAAD form

$$u(\mathbf{x},t)_t + \nabla\cdot((\mathbf{a}(\mathbf{x},t) + \mathbf{b}(\mathbf{x},t))u(t,\mathbf{x})) = \nabla\nabla\left(\mathbb{D}(\mathbf{x},t)u(\mathbf{x},t)\right)\,. \tag{48}$$

where $\mathbf{a}(\mathbf{x},t)$ and $\mathbb{D}(\mathbf{x},t)$ depend on the statistical inputs of the random walk (mean speed, turning rates, moments of the turning distribution).

Numerical methods for solving (48) are adapted from our previous studies (e.g. see [43]). We adopt a simple Method of Lines (MOL) approach, first discretising in space (using a fixed lattice of space $\Delta x$) to create a large system of ordinary differential equations (ODEs) which are subsequently integrated over time. The "fully anisotropic" diffusion term, is expanded into an advective and standard anisotropic-diffusion component. This advective component, along with advection terms arising from ocean currents and active directional swimming, are solved via a third-order upwinding scheme, augmented by flux-limiting to ensure positivity of solutions (e.g. see [24]). The choice of finite-difference discretisation for the anisotropic diffusion term is more specific: naive discretisations can lead to numerical instability for sufficiently anisotropic scenarios (high $\kappa$ values). The method of [58] allows greater flexibility in the choice of $\kappa$: in this scheme, finite difference derivatives are calculated and combined along distinct axial directions: the axes of the discretisation lattice and the major and minor axes of the ellipse corresponding to the anisotropic diffusion tensor. Under the moderate levels of anisotropy encountered here we obtain a stable scheme. Time discretisation here is performed via a simple forward Euler method with a suitably small time step.