MATHEMATICAL BIOSCIENCES
AND ENGINEERING
Volume **14**, Number **3**, June **2017**

doi:10.3934/mbe.2017038
pp. **673–694**

MOMENTS OF VON MISES AND FISHER DISTRIBUTIONS AND APPLICATIONS

Thomas Hillen

University of Alberta
Centre for Mathematical Biology
Edmonton, Alberta, T6G2G1, Canada

Kevin J. Painter

Department of Mathematics
Heriot-Watt University
Edinburgh, EH14 4AS, UK

Amanda C. Swan

University of Alberta
Centre for Mathematical Biology
Edmonton, Alberta, T6G2G1, Canada

Albert D. Murtha

Cross Cancer Institute
11560-University Ave NW
Edmonton, Alberta, T6G 1Z2, Canada

(Communicated by Hermann J. Eberl)

Abstract. The von Mises and Fisher distributions are spherical analogues to the Normal distribution on the unit circle and unit sphere, respectively. The computation of their moments, and in particular the second moment, usually involves solving tedious trigonometric integrals. Here we present a new method to compute the moments of spherical distributions, based on the divergence theorem. This method allows a clear derivation of the second moments and can be easily generalized to higher dimensions. In particular we note that, to our knowledge, the variance-covariance matrix of the three dimensional Fisher distribution has not previously been explicitly computed. While the emphasis of this paper lies in calculating the moments of spherical distributions, their usefulness is motivated by their relationship to population statistics in animal/cell movement models and demonstrated in applications to the modelling of sea turtle navigation, wolf movement and brain tumour growth.

1. **Introduction.** The motivation for the present paper stems from the modelling of oriented biological movement. In many studies of animal or cell movement, a primary point of concern is the cue (or cues) used to navigate, necessitating statistical correlation between a preferred movement direction and orientating information in the environment. Typical examples include the orientation of animals in response

2010 *Mathematics Subject Classification.* Primary: 35Q92; Secondary: 62E15, 92B05.
Key words and phrases. Von Mises distribution, Fisher distribution, spherical distributions, moments, biological applications.
* Corresponding author: Thomas Hillen.

to the earth's magnetic field [31], or the migration of invasive tumour cells through their local extracellular environment [30]. In ecological applications a quasi two-dimensional scenario generally suffices – for example, the habitat of a land-bound population or the typical swimming depth of marine organisms – and in such cases an animal's bearing can be described by a single angular coordinate. Other situations, however, may require a fully three-dimensional description, such as the motion of parasitic nematode worms in the soil, the movement of plankton in the oceans, or the flight of insects and birds. For the migration of cell populations in tissues a three dimensional understanding of movement can be particularly crucial. For example, malignant cancers of the central nervous system (CNS) are characterized by their diffuse and heterogeneous growth, and discovering the pathways of invasion is necessary for predicting tumour extent. The organisation of the brain itself plays a significant part in invasion, with cells channelled along the aligned neural fiber tracts that characterise white matter. Consequently, a three-dimensional understanding of the brain's architecture and the movement of cells in response to it can potentially lead to targeted treatment protocols.

To model these and other movement phenomena we will consider random walk models in the next section (Section 2). We will draw a connection between spherical distributions, also called *directional distributions*, and movement models for biological systems. We denote the spherical distribution by $q(\mathbf{n})$, where $\mathbf{n}$ is a unit vector indicating an individuals' bearing. It turns out that the first moment of $q(\mathbf{n})$ (i.e. the expectation $\mathbf{E}[q]$) is proportional to the population drift, while the variance-covariance matrix $\mathrm{Var}[q]$ is proportional to the diffusion of the population. Hence the moments of $q(\mathbf{n})$ provide immediate significance for interpreting population-level responses. In Section 3 we compute the first and second moments of the von Mises distribution in two dimensions. These moments are known, and we include our computation to introduce a method based on the divergence theorem. In Section 4 we apply this method to compute the first and second moment in three dimensions (in three dimensions the von Mises distribution is generally known as the Fisher or von Mises-Fisher distribution [20, 1][1]). To our knowledge, an explicit formula for the three dimensional variance is not known in the literature and here we provide such a formula. Moreover, this method generalises to higher space dimensions and other circular/spherical distributions. Finally, in Section 5 we illustrate the importance of these calculations to three applications of directional movement: the orientation of sea turtles in an ocean; the tendency of animals to follow or avoid linear landscape features such as roads or seismic lines; and, the three-dimensional anisotropic spread of glioma. In each of these cases we find that the von-Mises and/or Fisher distribution lies at the heart of the organizing centre for directed movement and that their moments provide key information on population-level distributions.

2. The von Mises and Fisher distributions.

We consider distributions $q(\mathbf{n})$ on the unit circle $\mathbb{S}^{n-1}$ in $\mathbb{R}^n$. The von Mises, Fisher and n-dimensional von Mises-Fisher distributions relate to a given direction $\mathbf{u} \in \mathbb{S}^{n-1}$, which identifies a "preferred" direction. In two dimensions the **von Mises distribution** is given as

$$q(\mathbf{n}) = \frac{1}{2\pi I_0(k)} e^{k\mathbf{n}\cdot\mathbf{u}}, \tag{1}$$

[1]Not to be confused with the statistical F-distribution.

for $\mathbf{n} \in \mathbb{S}^1$. Here k is the *concentration parameter* and $I_0(k)$, and later $I_j(k)$, denote the modified Bessel functions of first kind of order j. A more familiar representation of the von-Mises distribution in two dimensions uses the polar angle of the vectors involved. If we represent $\mathbf{n} = (\cos\alpha, \sin\alpha)^T$ and $\mathbf{u} = (\cos\mu, \sin\mu)^T$, then the two-dimensional von-Mises distribution is

$$q(\mathbf{n}) = \frac{1}{2\pi I_0(k)} e^{k\cos(\alpha-\mu)}.$$

We prefer the coordinate free vector notation in (1) as it generalizes easily and it is not coupled to a specific coordinate system. The same distribution in three dimensions is called the **Fisher distribution** and is given by

$$q(\mathbf{n}) = \frac{k}{4\pi\sinh(k)} e^{k\mathbf{n}\cdot\mathbf{u}}, \qquad \mathbf{n} \in \mathbb{S}^2. \tag{2}$$

The extension to higher dimensions forms the n-dimensional von Mises-Fisher distribution, given as

$$q(\mathbf{n}) = \frac{k^{n/2-1}}{(2\pi)^{n/2} I_{n/2-1}(k)} e^{k\mathbf{n}\cdot\mathbf{u}} \qquad \mathbf{n} \in \mathbb{S}^{n-1}.$$

The von Mises distribution occupies a predominant position in spherical statistics. As a form of 'wrapped' normal distribution it acts as the *de facto* standard for fitting oriented data and has received a great deal of attention (e.g. see [22]). Further, applying a linear combination (or mixture) of von Mises distributions can be used to fit more complicated datasets, such as skewed or multi-modal distributions, thereby extending its general utility. For a von Mises mixture with given unit vectors $\mathbf{u}^1, \ldots, \mathbf{u}^N$, we consider

$$q(\mathbf{n}) = \sum_{i=1}^{N} \frac{c_i}{2\pi I_0(k_i)} e^{k_i \mathbf{n}\cdot\mathbf{u}_i}, \tag{3}$$

where the coefficients c_i satisfy $\sum_{i=1}^N c_i = 1$. Of particular note is the symmetric, **bimodal von Mises distribution**

$$q_B(\mathbf{n}) = \frac{1}{4\pi I_0(k)} \left(e^{k\mathbf{n}\cdot\mathbf{u}} + e^{-k\mathbf{n}\cdot\mathbf{u}} \right), \tag{4}$$

along with its three-dimensional **bimodal Fisher** version

$$q(\mathbf{n}) = \frac{k}{8\pi\sinh(k)} \left(e^{k\mathbf{n}\cdot\mathbf{u}} + e^{-k\mathbf{n}\cdot\mathbf{u}} \right), \tag{5}$$

which define bidirectional distributions with two opposite and equal peaks. Figure 1 illustrates the standard von Mises distribution superimposed onto the unit circle, along with a bimodal von Mises distribution with peaks at $(1,0)^T$ and $(-1,0)^T$. Figure 2 illustrates the three dimensional distribution (2) for increasing values of k.

Given the prominent role of von Mises-Fisher distributions, we illustrate a novel method for computing their first and second moments. In the context of the application to animal/cell movement, this allows a direct line to be drawn between the statistical analysis of an experimentally-derived dataset and inputs for the predictive macroscopic-scale model (6). Typically, moments are derived via explicit trigonometric integration (e.g. [24, 3]) whereas our method here utilises the divergence theorem: while the former calculations become increasingly cumbersome as the space dimension increases, the latter generalises relatively easily.

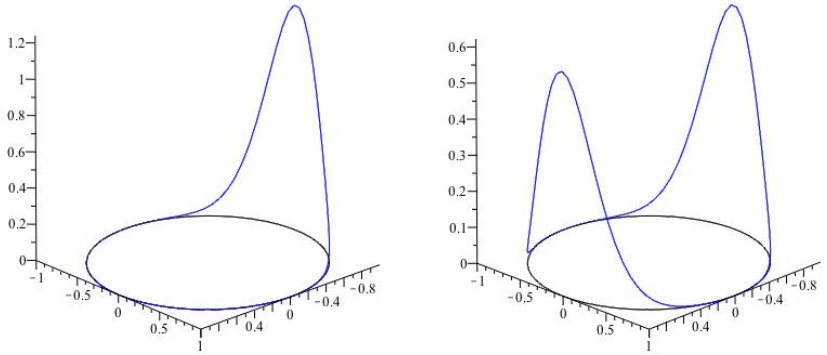


FIGURE 1. Left: The unimodal von Mises distribution as a function of $\mathbf{n} \in \mathbb{S}^1$ with a peak at $(1,0)^T$. Right: The bimodal von Mises distribution as a function of $\mathbf{n} \in \mathbb{S}^1$ with peaks at $\mathbf{n} = \pm(1,0)^T$.

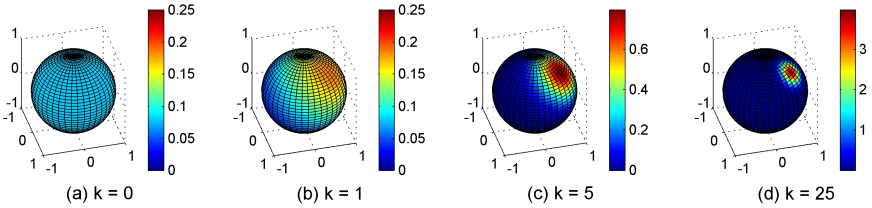


FIGURE 2. The Fisher distribution on the unit sphere. From left to right we plot the distribution for increasing values of concentration parameter k. In each case, the dominant direction is taken as $\mathbf{u} = (1/2, 1/2, \sqrt{2}/2)$.

To explain the use of the von Mises and Fisher distribution in population models, we consider a simple random walk model for oriented animal/cell movement paths in two and three dimensions. Random walks have an established history in the quantification and modelling of biological movement (e.g. see [27, 37, 7, 6, 38]). In the velocity-jump random walk, [28], a path is traced out as a series of 'runs' (continuous movement through space with constant velocity) punctuated with 'turns' (selecting a new velocity). The run and tumbling motion of flagellated bacteria such as *E. coli* provides the exemplar, yet more generally, it offers a natural description for movement, compatible with established methods of recording animal/cell tracks: datasets of turning angles, times *etc.* inform probability distributions for input into the velocity-jump model.

In the formulation here (see [14, 29, 16, 32] for further details and applications) we presume an individual (taken as a point) is characterized by its position ($\mathbf{x} \in \mathbb{R}^n$ for $n = 2, 3$) and velocity ($\mathbf{v} \in \mathbb{R}^n$) at time $t \in [0, \infty)$. For simplicity we assume the individual moves with a constant speed s, but that it undergoes changes in its direction ($\mathbf{n} \in \mathbb{S}^{n-1}$) as a result of a Poisson process, where λ is the (assumed constant) *turning rate* parameter. At each turn, a new direction ($\mathbf{n} \in \mathbb{S}^{n-1}$) is selected according to a suitable directional distribution $q(t, \mathbf{x}, \mathbf{n})$. Then $q(t, \mathbf{x}, \mathbf{n})$

will satisfy

$$\int_{\mathbb{S}^{n-1}} q(t,\mathbf{x},\mathbf{n})d\mathbf{n} = 1\,.$$

To obtain an efficient evolution equation for the macroscopic population density distribution $c(\mathbf{x},t)$, scaling techniques can be employed: we refer to [16] for full details. Adopting a moment-closure approach yields a macroscopic drift-diffusion equation in the form

$$\begin{aligned} c(t,\mathbf{x})_t + \nabla\cdot\Big(\mathbf{a}(t,\mathbf{x})c(t,\mathbf{x})\Big) &= \nabla\nabla : \Big(D(t,\mathbf{x})c(t,\mathbf{x})\Big)\,, \\ &= \sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}(D^{ij}c)\,, \end{aligned} \tag{6}$$

where the colon symbol (:) is used to denote the contraction of two tensors. In the above $D(t,\mathbf{x})$ defines a (symmetric) $n\times n$–diffusion tensor and $\mathbf{a}(t,\mathbf{x})$ is the $n-$dimensional drift (or advective) velocity. Formally, equation (6) describes the probability distribution for finding an individual at position $\mathbf{x}$ and time t; however, if the population is dispersed and interactions are negligible, it can be interpreted as a macroscopic population density distribution. Significantly, the macroscopic quantities $\mathbf{a}$ and D relate to the first two moments of the turning distribution q:

- the drift $\mathbf{a}$ is proportional to the *expectation* of q,

$$\mathbf{a}(t,\mathbf{x}) = s\mathbf{E}[q] = s\int_{\mathbb{S}^{n-1}} \mathbf{n}q(t,\mathbf{x},\mathbf{n})d\mathbf{n}\,; \tag{7}$$

- the D is proportional to the *variance-covariance matrix* of q,

$$D(t,\mathbf{x}) = \frac{s^2}{\lambda}\mathrm{Var}[q] = \frac{s^2}{\lambda}\int_{\mathbb{S}^{n-1}} (\mathbf{n}-\mathbf{a}(t,\mathbf{x})/s)(\mathbf{n}-\mathbf{a}(t,\mathbf{x})/s)^T q(t,\mathbf{x},\mathbf{n})d\mathbf{n}\,. \tag{8}$$

Note that the product $\mathbf{v}\mathbf{v}^T$ denotes the dyadic product of two vectors (in tensor notation: $\mathbf{v}\otimes\mathbf{v}$) and therefore D forms a symmetric matrix.

For the case in which q is unchanged across $\mathbf{x}$ and t, $\mathbf{a}$ and D will be given by a constant vector and diffusion tensor respectively. In this case we can calculate the fundamental solution of (6) as

$$c(\mathbf{x},t) = \frac{1}{\sqrt{\det D(4\pi t)^n}}\exp\left(-\frac{1}{4t}(\mathbf{x}-\mathbf{a}t)^T D^{-1}(\mathbf{x}-\mathbf{a}t)\right). \tag{9}$$

The above describes the probability density for finding a random walker (starting at location $\mathbf{0}$, time 0) at location $\mathbf{x}$, time t.

Clearly, the dynamics predicted by (6) pivot about the directional distribution q which, in turn, is dictated by data. Ideally, an explicit form for the moments of q can be calculated, allowing a direct translation between the inputs of the stochastic/individual model and the macroscopic model. Yet, while the statistical methodology for analysing oriented datasets has grown substantially in recent decades (e.g. see [3, 10, 22]), it remains substantially less advanced than that of classical (linear) statistics.

3. Moments of the von Mises distribution.

Lemma 3.1. *For a given unit vector $\mathbf{u}\in\mathbb{S}^1$ consider the von-Mises distribution*

$$q(\mathbf{n}) = \frac{1}{2\pi I_0(k)}e^{k\mathbf{n}\cdot\mathbf{u}}, \qquad \mathbf{n}\in\mathbb{S}^1.$$

Then q has the following expectation and variance:

$$\mathbf{E}[q] = \frac{I_1(k)}{I_0(k)}\,\mathbf{u}\,; \tag{10}$$

$$\mathrm{Var}[q] = \frac{1}{2}\left(1-\frac{I_2(k)}{I_0(k)}\right)\mathbb{I}_2 + \left(\frac{I_2(k)}{I_0(k)} - \left(\frac{I_1(k)}{I_0(k)}\right)^2\right)\mathbf{u}\mathbf{u}^T\,. \tag{11}$$

In the above, I_j denote modified Bessel functions of first kind and order j, $\mathbb{I}_2$ denotes the 2×2 identity matrix, and $\mathbf{u}\mathbf{u}^T$ denotes the dyadic product of two vectors (in tensor notation $\mathbf{u}\otimes\mathbf{u}$).

Proof [2]. Before we compute the moments of $q(\mathbf{n})$ we collect some properties of Bessel functions (see e.g. [15]). Let $J_n(x)$ denote the Bessel functions of first kind, then

$$I_n(x) := (-i)^{-n}J_n(ix)$$

denote Bessel functions of first kind with purely imaginary argument, or the *modified Bessel functions*. The modified Bessel functions have an explicit representation as

$$I_n(k) = \frac{1}{2\pi}\int_0^{2\pi}\cos(n\phi)e^{k\cos\phi}d\phi\,. \tag{12}$$

Bessel functions satisfy a differential relation

$$\frac{d}{dx}\left(x^nJ_n(x)\right) = x^nJ_{n-1}(x) \tag{13}$$

and, for $n\geq 1$, they satisfy a recurrence relation

$$J_{n+1}(x) = \frac{2n}{x}J_n(x) - J_{n-1}(x). \tag{14}$$

The total mass of the unimodal von Mises distribution (1) can be directly computed from (12), where we denote the angle between $\mathbf{n}$ and $\mathbf{u}$ as ϕ:

$$\int_{\mathbb{S}^1} q(\mathbf{n})d\mathbf{n} = \frac{1}{2\pi I_0(k)}\int_0^{2\pi}e^{k\cos\phi}d\phi = 1.$$

For the expectation,

$$\mathbf{E}[q] = \int_{\mathbb{S}^1}\mathbf{n}q(\mathbf{n})d\mathbf{n}\,,$$

we use a test vector $\mathbf{b}\in\mathbb{R}^2$ and multiply it by $\mathbf{E}$[q] (note that the expectation here is a vector quantity), using summation convention for repeated indices. Then we have

$$\mathbf{b}\cdot\mathbf{E}[q] = \mathbf{b}\cdot\int_{\mathbb{S}^1}\mathbf{n}q(\mathbf{n})d\mathbf{n},$$

and

$$\begin{aligned}\mathbf{b}\cdot 2\pi I_0(k)E[q] &= \mathbf{b}\cdot\int_{\mathbb{S}^1}\mathbf{n}e^{k\mathbf{n}\cdot\mathbf{u}}d\mathbf{n}\,,\\ &= \int_{\mathbb{S}^1}\mathbf{n}^i(\mathbf{b}_ie^{k\mathbf{n}\cdot\mathbf{u}})d\mathbf{n}\,,\end{aligned}$$

[2]The computations in this section have already been published in an appendix to a book chapter [16]. We include them here for completeness and to introduce our method.

$$
\begin{aligned}
&= \int_{\mathbb{B}_1(0)} \frac{\partial}{\partial \mathbf{v}^i}(\mathbf{b}_i e^{k\mathbf{v}\cdot\mathbf{u}}) d\mathbf{v}\,, \\
&= \int_{\mathbb{B}_1(0)} \mathbf{b}_i k \mathbf{u}_i e^{k\mathbf{v}\cdot\mathbf{u}} d\mathbf{v}\,, \\
&= \mathbf{b}\cdot k\mathbf{u} \int_0^1 \int_0^{2\pi} e^{rk cos\phi} r dr d\phi\,, \\
&= \mathbf{b}\cdot k\mathbf{u} \int_0^1 2\pi r I_0(rk) dr\,, \\
&= \mathbf{b}\cdot 2\pi k\mathbf{u} \int_0^1 r I_0(rk) dr\,.
\end{aligned}
$$

In the first step we use the summation convention, which requires summation over repeated indices, i.e. $a^i b_i = \sum_{i=1}^n a^i b_i$. In the third step we used the divergence theorem to transform a surface integral over $\mathbb{S}^1$ into a volume integral over the unit ball $\mathbb{B}_1(0)$ of a divergence-term.

The last integral can be solved via (13) and

$$
r I_0(rk) = \frac{irk J_0(irk)}{ik} = \frac{1}{ik}\frac{d}{dx}(xJ_1(x))|_{x=irk} = \frac{1}{ik}\frac{d}{dr}(rJ_1(irk)).
$$

Then

$$
\int_0^1 r I_0(rk) dr = \frac{1}{ik} J_1(ik) = \frac{1}{ik} i I_1(k) = \frac{I_1(k)}{k}, \tag{15}
$$

and due to the fact that $\mathbf{b}$ is an arbitrary test vector

$$
\mathbf{E}[q] = \frac{I_1(k)}{I_0(k)}\,\mathbf{u},
$$

which is (10).

The variance-covariance matrix of q is given by

$$
\mathrm{Var}[q] = \int_{\mathbb{S}^1} (\mathbf{n} - \mathbf{E}[q])(\mathbf{n} - \mathbf{E}[q])^T q(\mathbf{n}) d\mathbf{n} = \int_{\mathbb{S}^1} \mathbf{n}\mathbf{n}^T q(\mathbf{n}) d\mathbf{n} - \mathbf{E}[q]\mathbf{E}[q]^T.
$$

To find the second moment of q we consider two test vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ and we use index notation for automatic summation over repeated indices:

$$
\begin{aligned}
& 2\pi I_0(k)\mathbf{a} \int_{\mathbb{S}^1} \mathbf{n}\mathbf{n}^T q(\mathbf{n}) d\mathbf{n}\ \mathbf{b} \qquad (16) \\
&= \int_{\mathbb{S}^1} \mathbf{a}_i \mathbf{n}^i \mathbf{b}_j \mathbf{n}^j e^{k\mathbf{n}^l \mathbf{u}_l} d\mathbf{n}, \\
&= \int_{\mathbb{S}^1} \mathbf{n}^i (\mathbf{a}_i \mathbf{b}_j \mathbf{n}^j e^{k\mathbf{n}^l \mathbf{u}_l}) d\mathbf{n}, \\
&= \int_{\mathbb{B}_1(0)} \frac{\partial}{\partial \mathbf{v}^i}(\mathbf{a}_i \mathbf{b}_j \mathbf{v}^j e^{k\mathbf{v}^l \mathbf{u}_l}) d\mathbf{v}, \\
&= \int_{\mathbb{B}_1(0)} \mathbf{a}_i \mathbf{b}_i e^{k\mathbf{v}\cdot\mathbf{u}} d\mathbf{v} + \int_{\mathbb{B}_1(0)} \mathbf{a}_i (\mathbf{v}\cdot\mathbf{b}) k\mathbf{u}_i e^{k\mathbf{v}\cdot\mathbf{u}} d\mathbf{v}, \\
&= \mathbf{a}\cdot\mathbf{b} \int_{\mathbb{B}_1(0)} e^{k\mathbf{v}\cdot\mathbf{u}} d\mathbf{v} + k\mathbf{a}\cdot\mathbf{u}\ \mathbf{b}\cdot \int_{\mathbb{B}_1(0)} \mathbf{v} e^{k\mathbf{v}\cdot\mathbf{u}} d\mathbf{v}, \qquad (17)
\end{aligned}
$$

where, again, we use the summation convention and the divergence theorem. The first integral in (17) can be solved directly:

$$\int_{\mathbb{B}_1(0)} e^{k\mathbf{v}\cdot\mathbf{u}} d\mathbf{v} = \int_0^1 \int_{\mathbb{S}^1} e^{rk\mathbf{n}\cdot\mathbf{u}} r dr d\mathbf{n} = \int_0^1 2\pi r I_0(rk) dr = 2\pi \frac{I_1(k)}{k},$$

where we used (12) and (15) in the penultimate and ultimate step, respectively. Using (12) we can transform the second integral from (17) as follows:

$$\begin{aligned} \int_{\mathbb{B}_1(0)} \mathbf{v} e^{k\mathbf{v}\cdot\mathbf{u}} d\mathbf{v} &= \int_0^1 \int_{\mathbb{S}^1} r\mathbf{n} e^{rk\mathbf{n}\cdot\mathbf{u}} r dr d\mathbf{n} = \int_0^1 r^2 \int_{\mathbb{S}^1} \mathbf{n} e^{rk\mathbf{n}\cdot\mathbf{u}} d\mathbf{n} dr, \\ &= 2\pi\mathbf{u} \int_0^1 r^2 I_1(rk) dr, \end{aligned} \tag{18}$$

where we used (10) in the last step.

Now we use the differential recurrence relation (13) to write

$$r^2 I_1(rk) = -\frac{1}{ik^2}(irk)^2 J_1(irk) = -\frac{1}{ik^2}\frac{d}{dx}(x^2 J_1(x))|_{x=irk} = -\frac{1}{k}\frac{d}{dr}(r^2 J_1(irk)).$$

Continuing from (18) we find

$$\int_{\mathbb{B}_1(0)} \mathbf{v} e^{k\mathbf{v}\cdot\mathbf{u}} d\mathbf{v} = -2\pi\mathbf{u} \int_0^1 \frac{1}{k}\frac{d}{dr}(r^2 J_1(irk)) dr = -2\pi\mathbf{u} J_2(ik) = 2\pi\mathbf{u}\frac{I_2(k)}{k}. \tag{19}$$

Substituting all the integrals back into equation (17):

$$\begin{aligned} \mathbf{a} \int_{\mathbb{S}^1} \mathbf{n}\mathbf{n}^T q(\mathbf{n}) d\mathbf{n}\, \mathbf{b} &= \mathbf{a}\cdot\mathbf{b}\frac{2\pi\frac{I_1(k)}{k}}{2\pi I_0(k)} + k\mathbf{a}\cdot\mathbf{u}\frac{2\pi\mathbf{u}\cdot\mathbf{b}\frac{I_2(k)}{k}}{2\pi I_0(k)}, \\ &= \mathbf{a}\left(\frac{1}{k}\frac{I_1(k)}{I_0(k)}\mathbb{I}_2 + \mathbf{u}\mathbf{u}^T\frac{I_2(k)}{I_0(k)}\right)\mathbf{b}. \end{aligned}$$

Finally, we use the identity (14) for $n = 1$ to replace

$$\frac{1}{k}\frac{I_1(k)}{I_0(k)} = \frac{1}{2}\left(1 - \frac{I_2(k)}{I_0(k)}\right),$$

and the second moment is given by

$$\int_{\mathbb{S}^1} \mathbf{n}\mathbf{n}^T q(\mathbf{n}) d\mathbf{n} = \frac{1}{2}\mathbb{I}_2 + \frac{I_2(k)}{I_0(k)}\left(\mathbf{u}\mathbf{u}^T - \frac{1}{2}\mathbb{I}_2\right). \tag{20}$$

Together with the formula for the expectation (10) we find

$$\begin{aligned} \mathrm{Var}[q] &= \int_{\mathbb{S}^1} \mathbf{n}\mathbf{n}^T q(\mathbf{n}) d\mathbf{n} - \mathbf{E}[q]\mathbf{E}[q]^T, \\ &= \frac{1}{2}\mathbb{I}_2 + \frac{I_2(k)}{I_0(k)}\left(\mathbf{u}\mathbf{u}^T - \frac{1}{2}\mathbb{I}_2\right) - \left(\frac{I_1(k)}{I_0(k)}\right)^2 \mathbf{u}\mathbf{u}^T, \\ &= \frac{1}{2}\left(1 - \frac{I_2(k)}{I_0(k)}\right)\mathbb{I}_2 + \left(\frac{I_2(k)}{I_0(k)} - \left(\frac{I_1(k)}{I_0(k)}\right)^2\right)\mathbf{u}\mathbf{u}^T, \end{aligned}$$

which is (11). □

In the two dimensional case, it is useful to formulate the result of Lemma 10 in terms of a polar angle. For $\mathbf{u} = (\cos\mu, \sin\mu)^T$ we find

$$\begin{aligned} \mathbf{E}[q] &= \frac{I_1(k)}{I_0(k)} \begin{pmatrix} \cos\mu \\ \sin\mu \end{pmatrix}; & (21) \\ \mathrm{Var}[q] &= \frac{1}{2}\left(1 - \frac{I_2(k)}{I_0(k)}\right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &\quad + \left(\frac{I_2(k)}{I_0(k)} - \left(\frac{I_1(k)}{I_0(k)}\right)^2\right) \begin{pmatrix} \cos^2\mu & \cos\mu\sin\mu \\ \cos\mu\sin\mu & \sin^2\mu \end{pmatrix}. & (22) \end{aligned}$$

As outlined earlier, for many applications it is useful to compute the first and second moments for the bimodal von Mises distribution.

Corollary 1. *For given $\mathbf{u} \in \mathbb{S}^1$ consider the bimodal von Mises distribution*

$$q(\mathbf{n}) = \frac{1}{4\pi I_0(k)} \left(e^{k\mathbf{n}\cdot\mathbf{u}} + e^{-k\mathbf{n}\cdot\mathbf{u}}\right),$$

then

$$\begin{aligned} \mathbf{E}[q] &= 0, & (23) \\ \mathrm{Var}[q] &= \frac{1}{2}\left(1 - \frac{I_2(k)}{I_0(k)}\right)\mathbb{I}_2 + \frac{I_2(k)}{I_0(k)}\mathbf{u}\mathbf{u}^T. & (24) \end{aligned}$$

Proof. Since the bimodal von Mises distribution is symmetric we have $\mathbf{E}[q] = 0$ and $\mathrm{Var}[q] = \int \mathbf{n}\mathbf{n}^T q(\mathbf{n}) d\mathbf{n}$. We apply formula (20) for each of the components $e^{k\mathbf{n}\cdot\mathbf{u}}$ and $e^{-k\mathbf{n}\cdot\mathbf{u}}$ separately and sum. □

It is interesting to consider the limiting scenario as the parameter of concentration k becomes small (i.e. $k \to 0$). Here, the directional bias becomes increasingly small and (1) or (4) limit to a uniform distribution. A straightforward exploration reveals that $I_1(k)/I_0(k), I_2(k)/I_0(k) \to 0$. Consequently, $\mathbf{E}[q] \to 0$ while the variance-covariance $\mathrm{Var}[q] \to \frac{1}{2}\mathbb{I}_2$ for both the unimodal and bimodal cases.

4. The 3-D Fisher distribution.

The main result of this paper is as follows.

Theorem 4.1. *For given $\mathbf{u} \in \mathbb{S}^2$ consider the Fisher distribution*

$$q(\mathbf{n}) = \frac{k}{4\pi \sinh k} e^{k\mathbf{u}\cdot\mathbf{n}}, \qquad \mathbf{n} \in \mathbb{S}^2.$$

Then

$$\begin{aligned} \mathbf{E}[q] &= \left(\coth k - \frac{1}{k}\right)\mathbf{u}, & (25) \\ Var[q] &= \left(\frac{\coth k}{k} - \frac{1}{k^2}\right)\mathbb{I} + \left(1 - \frac{\coth k}{k} + \frac{2}{k^2} - \coth^2 k\right)\mathbf{u}\mathbf{u}^T. & (26) \end{aligned}$$

Proof. To compute the mean direction

$$\mathbf{E}[q] = \int_{\mathbb{S}^2} \mathbf{n} q(\mathbf{n}) d\mathbf{n},$$

we use a test-vector $\mathbf{b} \in \mathbb{R}^3$ and multiply it by $\mathbf{E}[q]$. Using summation convention for repeated indices:

$$\mathbf{b}\cdot\mathbf{E}[q] = \mathbf{b}\cdot\int_{\mathbb{S}^2} \mathbf{n} q(\mathbf{n}) d\mathbf{n},$$

$$
\begin{aligned}
&= \int_{\mathbb{S}^2} \mathbf{n}^i (\mathbf{b}_i q(\mathbf{n})) d\mathbf{n}\,, \\
&= \int_{\mathbb{B}_1(0)} \frac{\partial}{\partial \mathbf{v}^i} \left(\mathbf{b}_i \frac{k}{4\pi \sinh k} e^{k \mathbf{u}_l \mathbf{v}^l} \right) d\mathbf{v}\,, \\
&= \mathbf{b}^i \int_{\mathbb{B}_1(0)} \frac{k^2 \mathbf{u}_i}{4\pi \sinh k} e^{k\mathbf{u}\cdot\mathbf{v}} d\mathbf{v}\,, \\
&= \mathbf{b}\cdot\mathbf{u} k \int_{\mathbb{B}_1(0)} \frac{k}{4\pi \sinh k} e^{k\mathbf{u}\cdot\mathbf{v}} d\mathbf{v}\,, \\
&= \mathbf{b}\cdot\mathbf{u} k \int_0^1 \int_{\mathbb{S}^2} \frac{k}{4\pi \sinh k} e^{rk\mathbf{u}\cdot\mathbf{n}} r^2 dr d\mathbf{n}\,, \\
&= \mathbf{b}\cdot\mathbf{u} k \int_0^1 \frac{r^2 k}{4\pi \sinh k} \frac{4\pi \sinh rk}{rk} dr\,, \\
&= \frac{\mathbf{b}\cdot\mathbf{u} k}{\sinh k} \left(-\frac{1}{k^2} \sinh k + \frac{1}{k} \cosh k \right)\,, \\
&= \mathbf{b}\cdot\mathbf{u} \left(\coth k - \frac{1}{k} \right),
\end{aligned}
$$

where in the 5th step we used the normalization condition for a Fisher distribution with concentration parameter rk, i.e.

$$\int_{\mathbb{S}^2} e^{rk\mathbf{u}\cdot\mathbf{n}} d\mathbf{n} = \frac{4\pi \sin(rk)}{rk}.$$

Since the test vector $\mathbf{b}$ was arbitrary, we proved (25). This formula coincides with formula (9.3.9) in Mardia [22].

To compute the second moment, we take two test-vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ and compute

$$
\begin{aligned}
&\mathbf{a} \cdot \int_{\mathbb{S}^2} \mathbf{n}\mathbf{n}^T q(\mathbf{n}) d\mathbf{n} \cdot \mathbf{b} \\
&= \int_{\mathbb{S}^2} \mathbf{a}_i \mathbf{n}^i \mathbf{n}^j \mathbf{b}_j q(\mathbf{n}) d\mathbf{n}\,, \\
&= \int_{\mathbb{S}^2} \mathbf{n}^i (\mathbf{a}_i \mathbf{n}^j \mathbf{b}_j q(\mathbf{n})) d\mathbf{n}\,, \\
&= \int_{\mathbb{B}_1(0)} \frac{\partial}{\partial \mathbf{v}^i} \left(\mathbf{a}_i \mathbf{v}^j \mathbf{b}_j \frac{k}{4\pi \sinh k} e^{k \mathbf{u}_l \mathbf{v}^l} \right) d\mathbf{v}\,, \\
&= \int_{\mathbb{B}_1(0)} \mathbf{a}_i \mathbf{b}_i \frac{k}{4\pi \sinh k} e^{k\mathbf{u}\cdot\mathbf{v}} d\mathbf{v} + \int_{\mathbb{B}_1(0)} \mathbf{a}_i (\mathbf{v}\cdot\mathbf{b}) \frac{k^2 \mathbf{u}^i}{4\pi \sinh k} e^{k\mathbf{u}\cdot\mathbf{v}} d\mathbf{v}\,, \\
&= \underbrace{\mathbf{a}\cdot\mathbf{b} \int_{\mathbb{B}_1(0)} \frac{k}{4\pi \sinh k} e^{k\mathbf{u}\cdot\mathbf{v}} d\mathbf{v}}_{(A)} + \underbrace{\mathbf{a}\cdot\mathbf{u}\ \mathbf{b}\cdot \int_{\mathbb{B}_1(0)} \frac{k^2}{4\pi \sinh k} \mathbf{v} e^{k\mathbf{u}\cdot\mathbf{v}} d\mathbf{v})}_{(B)}. \qquad (27)
\end{aligned}
$$

From the computation of the first moment, we find for the first integral (A) that

$$\int_{\mathbb{B}_1(0)} \frac{k}{4\pi \sinh k} e^{k\mathbf{u}\cdot\mathbf{v}} d\mathbf{v} = \frac{1}{k} \coth k - \frac{1}{k^2}. \qquad (28)$$

The second integral (B) can be computed as

$$\int_{\mathbb{B}_1(0)} \mathbf{v} e^{k\mathbf{u}\cdot\mathbf{v}} d\mathbf{v} = \int_0^1 \int_{\mathbb{S}^2} r^3 \mathbf{n} e^{rk\mathbf{u}\cdot\mathbf{n}} d\mathbf{n} dr.$$

From the first moment we obtain

$$\int_{\mathbb{S}^2} \mathbf{n} e^{rk\mathbf{u}\cdot\mathbf{n}} d\mathbf{n} = \left[\frac{4\pi \sinh rk}{rk}\left(\coth rk - \frac{1}{rk}\right)\right]\mathbf{u}.$$

Hence

$$\begin{aligned}\int_{\mathbb{B}_1(0)} \mathbf{v} e^{k\mathbf{u}\cdot\mathbf{v}} d\mathbf{v}\,, &= \left[\int_0^1 r^3 \frac{4\pi \sinh rk}{rk}\left(\coth rk - \frac{1}{rk}\right) dr\right]\mathbf{u}\,,\\ &= 4\pi\left[\int_0^1 \frac{r^2\cosh rk}{k} dr - \int_0^1 \frac{r\sinh rk}{k^2} dr\right]\mathbf{u}\,,\\ &= 4\pi\left[\frac{\sinh k}{k^2} - \frac{3\cosh k}{k^3} + \frac{3\sinh k}{k^4}\right]\mathbf{u}.\end{aligned}$$

Then

$$\frac{k^2}{4\pi\sinh k}\int_{\mathbb{B}_1(0)} \mathbf{v} e^{k\mathbf{u}\cdot\mathbf{v}} d\mathbf{v} = \left(1 - \frac{3}{k}\coth k + \frac{3}{k^2}\right)\mathbf{u}. \tag{29}$$

Combining (28) and (29) with (27) we obtain

$$\mathbf{a}\cdot\int \mathbf{n}\mathbf{n}^T q(\mathbf{n}) d\mathbf{n}\cdot\mathbf{b} = \mathbf{a}\cdot\mathbf{b}\left(\frac{\coth k}{k} - \frac{1}{k^2}\right) + \mathbf{a}\cdot\mathbf{u}\,\mathbf{u}\cdot\mathbf{b}\left(1 - \frac{3\coth k}{k} + \frac{3}{k^2}\right),$$

and the second moment is given by

$$\int \mathbf{n}\mathbf{n}^T q(\mathbf{n}) d\mathbf{n} = \left(\frac{\coth k}{k} - \frac{1}{k^2}\right)\mathbb{I} + \left(1 - \frac{3\coth k}{k} + \frac{3}{k^2}\right)\mathbf{u}\mathbf{u}^T. \tag{30}$$

To compute the full variance-covariance matrix of q we use the formula $\mathrm{Var}[q] = \int \mathbf{n}\mathbf{n}^T q(\mathbf{n}) d\mathbf{n} - \mathbf{E}[q]\mathbf{E}[q]^T$ to obtain

$$\mathrm{Var}[q] = \left(\frac{\coth k}{k} - \frac{1}{k^2}\right)\mathbb{I} + \left(1 - \frac{\coth k}{k} + \frac{2}{k^2} - \coth^2 k\right)\mathbf{u}\mathbf{u}^T,$$

which is (26). We were not able to find this formula elsewhere in the literature. □

Corollary 2. *The first two moments of the bimodal Fisher distribution*

$$q(\mathbf{n}) = \frac{k}{8\pi\sinh k}\left(e^{k\mathbf{u}\cdot\mathbf{n}} + e^{-k\mathbf{u}\cdot\mathbf{n}}\right) = \frac{k}{4\pi\sinh k}\cosh(k\mathbf{u}\cdot\mathbf{n})$$

are given by

$$\mathbf{E}[q] = 0, \tag{31}$$

$$\mathrm{Var}[q] = \left(\frac{\coth k}{k} - \frac{1}{k^2}\right)\mathbb{I} + \left(1 - \frac{3\coth k}{k} + \frac{3}{k^2}\right)\mathbf{u}\mathbf{u}^T. \tag{32}$$

Proof. As for the bimodal von Mises distribution, we exploit the symmetry properties by considering the components $e^{k\mathbf{n}\cdot\mathbf{u}}$ and $e^{-k\mathbf{n}\cdot\mathbf{u}}$ separately and summing. □

It can be revealing to consider the coefficients arising in (25, 26, 32) and their dependence on the concentration parameter k. Consider the following four coefficients as functions of k:

$$\begin{aligned} a(k) &= \coth k - \frac{1}{k}\,;\\ b(k) &= \frac{\coth k}{k} - \frac{1}{k^2}\,;\\ c(k) &= 1 - \frac{\coth k}{k} + \frac{2}{k^2} - \coth^2 k\,;\\ d(k) &= 1 - \frac{3\coth k}{k} + \frac{3}{k^2}. \end{aligned} \tag{33}$$

Then for the unimodal case

$$\mathbf{E}[q] = a(k)\mathbf{u}, \qquad \mathrm{Var}[q] = b(k)\mathbb{I} + c(k)\mathbf{u}\mathbf{u}^T, \tag{34}$$

and for the bimodal case

$$\mathbf{E}[q] = 0, \qquad \mathrm{Var}[q] = b(k)\mathbb{I} + d(k)\mathbf{u}\mathbf{u}^T. \tag{35}$$

Thus, the variance-covariance matrix is split into an isotropic component, determined by $b(k)$, and an anisotropic component, determined by $c(k)$ or $d(k)$, respectively. In two dimensions as the (unimodal and bimodal) von Mises distribution tends towards a uniform distribution ($k \to 0$) we found that the expectation vanishes while the variance-covariance becomes an isotropic matrix. While substituting $k = 0$ directly into (25, 26, 32) is not possible, the limit as $k \to 0$ for the various coefficients above can be computed. Using the above notation,

$$\lim_{k\to 0^+} a(k) = \lim_{k\to 0^+} \coth k - \frac{1}{k},$$

which can be rewritten using l'Hôpital's rule as

$$\begin{aligned}
\lim_{k\to 0^+} a(k) &= \lim_{k\to 0^+} \frac{k - \tanh k}{k \tanh k}, \\
&= \lim_{k\to 0^+} \frac{1 - \mathrm{sech}^2 k}{\tanh k + k\mathrm{sech}^2 k}, \\
&= \lim_{k\to 0^+} \frac{2\mathrm{sech}^2 k \tanh k}{\mathrm{sech}^2 k + (1 - 2\mathrm{sech} k)\mathrm{sech}^2 k}, \\
&= 0.
\end{aligned}$$

Continuing in a similar fashion for $b(k)$, we have

$$\lim_{k\to 0^+} b(k) = \lim_{k\to 0^+} \frac{\coth k}{k} - \frac{1}{k^2}.$$

This can be rewritten with repeated use of l'Hôpital's rule as

$$\begin{aligned}
\lim_{k\to 0^+} b(k) &= \lim_{k\to 0^+} \frac{k - \tanh k}{k^2 \tanh k}, \\
&= \lim_{k\to 0^+} \frac{1 - \mathrm{sech}^2 k}{2k \tanh k + k^2 \mathrm{sech}^2 k}, \\
&= \lim_{k\to 0^+} \frac{2\mathrm{sech}^2 k \tanh k}{2k\mathrm{sech}^2 k + 2\tanh k - 2k^2 \mathrm{sech}^2 k \tanh k + 2k\mathrm{sech}^2 k}, \\
&= \lim_{k\to 0^+} \frac{-4\mathrm{sech}^2 k \tanh^2 k + 2\mathrm{sech}^4 k}{6\mathrm{sech}^2 k - 12k\mathrm{sech}^2 k \tanh k - 2k^2(\mathrm{sech}^4 k - 2\mathrm{sech}^2 k \tanh^2 k)}, \\
&= \frac{1}{3}.
\end{aligned}$$

Then we have

$$\lim_{k\to 0^+} c(k) = \lim_{k\to 0^+} 1 - \frac{\coth k}{k} + \frac{2}{k^2} - \coth^2 k,$$

which we may write as

$$\lim_{k\to 0^+} 1 - \underbrace{\left(\frac{\coth k}{k} - \frac{1}{k^2}\right)}_{b(k)} + \frac{1}{k^2} - \coth^2 k = \lim_{k\to 0^+} \frac{2}{3} + \left(\frac{1}{k} + \coth k\right)\left(\frac{1}{k} - \coth k\right).$$

In the above we have used the fact that the limit of $b(k)$ is $1/3$. Taking a factor of $1/k$ from the first factor to the second, we have

$$\begin{aligned}\lim_{k\to 0^+} c(k) &= \lim_{k\to 0^+} \frac{2}{3} + \left(\frac{1}{k} + \coth k\right)\left(\frac{1}{k} - \coth k\right), \\ &= \lim_{k\to 0^+} \frac{2}{3} - (1 + k\coth k)\,(b(k)), \\ &= \lim_{k\to 0^+} \frac{2}{3} - \frac{1}{3}\left(1 + \frac{k}{\sinh k}\right), \\ &= \frac{2}{3} - \frac{2}{3} = 0,\end{aligned}$$

where we used l'Hôpital's rule in the last step. Finally, for $d(k)$ we have

$$\begin{aligned}\lim_{k\to 0^+} d(k) &= \lim_{k\to 0^+} 1 - \frac{3\coth k}{k} + \frac{3}{k^2}, \\ &= \lim_{k\to 0^+} 1 - 3\left(\frac{\coth k}{k} - \frac{1}{k^2}\right), \\ &= 1 - 3\lim_{k\to 0^+} b(k), \\ &= 1 - 3\left(\frac{1}{3}\right) = 0.\end{aligned}$$

Summarising, for either the unimodal or bimodal Fisher distribution, as $k \to 0$ we find

$$\mathbf{a} = \mathbf{0} \quad \text{and} \quad D = \tau s^2 \frac{1}{3}\mathbb{I}_3$$

as would be expected.

These results are demonstrated numerically by plotting the coefficients as a function of k, see Figure 3: as k decreases to 0, the directional distribution tends to the uniform distribution and the variance-covariance matrix for either the unimodal or bimodal case limits to an isotropic tensor, as $c(k)$ and $d(k)$ tend to 0. Furthermore, we can clearly see from these plots that as anisotropy increases (k increases), the expectation of the uni-modal distribution becomes more strongly aligned with $\mathbf{u}$.

As a final note of interest, it is remarked that for the unimodal case (34) the coefficient $c(k)$ that describes the anisotropic diffusion component is negative. This effectively means reduced diffusion in the dominant direction $\mathbf{u}$: the unimodal von-Mises distribution describes enhanced movement in direction $\mathbf{u}$, with a correspondingly reduced movement in direction $-\mathbf{u}$, thereby reducing the variance in this direction.

5. **Numerical tests and applications.** We use this section to briefly highlight three interesting applications. In the interest of brevity we refer to the literature for biological and modelling details. We first compare our macroscopic PDE model (6) to an explicit random walk simulation in Section 5.1. The results show excellent overlap between the continuous model and the Monte Carlo simulation. In Section 5.2 we consider orientation of sea turtles as described by the unimodal two dimensional von Mises distribution. Section 5.3 considers the use of a bimodal two-dimensional von Mises distribution for animal movement patterns in disturbed forest environments. Finally, in Section 5.4, we use a bimodal three-dimensional Fisher distribution to model brain tumour growth. In all cases, simulations of

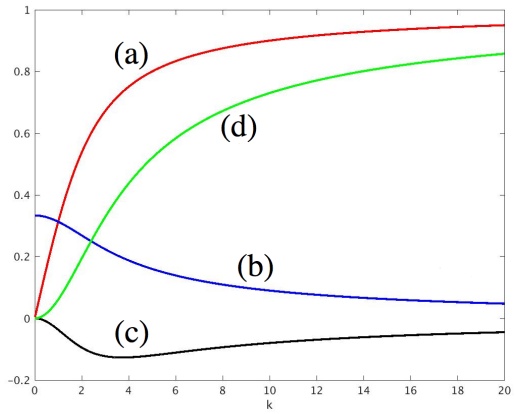


FIGURE 3. Coefficients $a(k), b(k), c(k), d(k)$ from (33) as functions of k

the population-level animal/cell density distributions exploit the explicit moment calculations derived in Sections 3 and 4.

5.1. **Fixed, homogeneous orienteering cue.** To illustrate anisotropic diffusion we consider the movement of individuals released at $\mathbf{x} = \mathbf{0}$ at $t = 0$ and subjected to a fixed, spatially uniform orientating cue. Consequently, q is fixed with respect to $\mathbf{x}$ and t and the macroscopic density distribution predicted by (6) will be determined by the fundamental solution (9), with $\mathbf{a}$ and D given by (7)–(8).

We initially consider a 2D scenario, where q is given by a von Mises distribution with $\phi = \pi/4$ and $\kappa = 10$. In Figure 4 the density distribution for particle position is generated through repeated (100000) simulations of the individual model and compared against the fundamental solution (9) with $\mathbf{a}$ and D determined by equations (10) and (11). The correspondence between solutions validates both the derivation of (6) as well as the calculations of $\mathbf{a}$ and D. In Figure 5 we show a similar comparison for a 3-dimensional scenario, using the Fisher distribution (2) and correspondingly the moments (25–26) for $\mathbf{a}$ and D.

5.2. **Sea turtle navigation.** The ability of animals to navigate has fascinated scientists for centuries. Certain green sea turtles (*Chelonia mydas*) are particularly renowned for swimming halfway across the Atlantic Ocean to nest at tiny Ascension Island [25], a feat noted already by Charles Darwin [8]. While the precise navigating cues remain uncertain, a combination of factors from geomagnetic information to wind or water-borne odours have been proposed to help turtles on their journey. In [31] we formulated an individual-based model for turtle navigation, where the turning direction for a simulated turtle from its current location was biased towards the island via a two dimensional von Mises distribution. The subsequent population-level macroscopic model is given by equation (6), although augmented by an additional advection term to account for ocean current drift, where the first two moments from equation (1) provide the forms for $\mathbf{a}$ and D. We refer to [31] for full details of the modelling.

Figure 6 shows the results from two simulation scenarios. All simulations consider a start date of 1st of January 2014 (early in a typical nesting season) and

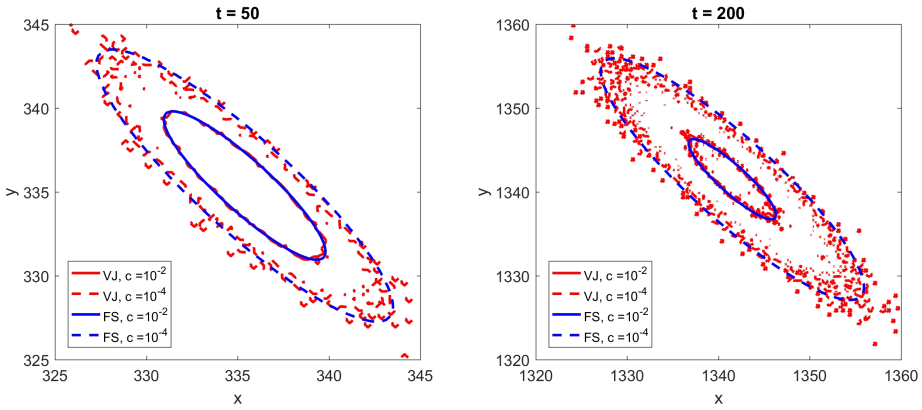


FIGURE 4. Comparison between simulations of the stochastic velocity-jump model and the fundamental solution of (6), equation (9) in two dimensions. Results shown at times $t = 50$ (left) and $t = 200$ (right). Red lines indicate contours of the average distribution for the particle position, generated through repeated simulations of the stochastic velocity-jump (VJ) model, while blue lines indicate contours of the fundamental solution (FS), equation (9); contours are shown here for $c(\mathbf{x}, t) = 10^{-4}$ and 10^{-2}. For the stochastic simulations we initialize the individual at $(0, 0)$ and compute the random walk path using the von Mises distribution for q (with $\phi = \pi/4$ and $\kappa = 10$), $s = 10$ and $\lambda = 100$. $\mathbf{a}$ and D for (9) are subsequently determined using (10) and (11).

a population of 1000 turtles that are initially distributed uniformly throughout a circular region centred on the island. Simulated turtles subsequently attempt to swim towards the island, with some fixed concentration parameter k that reflects the "navigational strength" of the cue being utilised. For these simulations, the additional drift due to ocean currents is taken from (post-validated) ocean surface flow data provided by an ocean forecasting model [5] over the studied time period. Figure 6 shows simulations for populations of (a) "weak navigators" ($k = 0.5$) and (b) "strong navigators" ($k = 2.0$). For each case we show results from (top panel) the underlying individual-based model and (bottom panel) the macroscopic model (6) for the population distribution: the close match between these distributions demonstrates that the macroscopic model (6) can accurately capture the individual-level data and stresses the usefulness of the moment calculations. Our simulations show that weak-navigators are generally unsuccessful in their attempts to reach the island: the orientating bias is not sufficiently strong to overcome the often powerful ocean currents encountered. Strong navigators, however, are able to plot a relatively straightforward course to the island. More details, in particular on the estimation of model parameters, can be found in [31].

5.3. **Wolf movement on seismic lines.** A good example in the usage of the bimodal von Mises distribution comes from wolf movement in disturbed forest environments. Boreal landscapes in Northern Alberta are characterized by a network of narrow straight corridors of cleared forest: so-called "seismic lines" created by

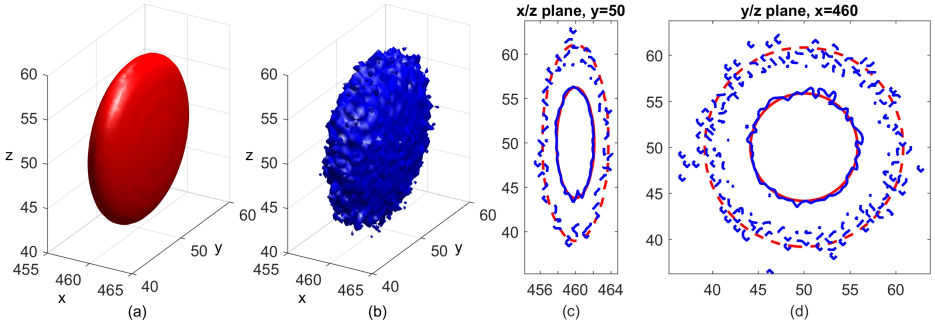
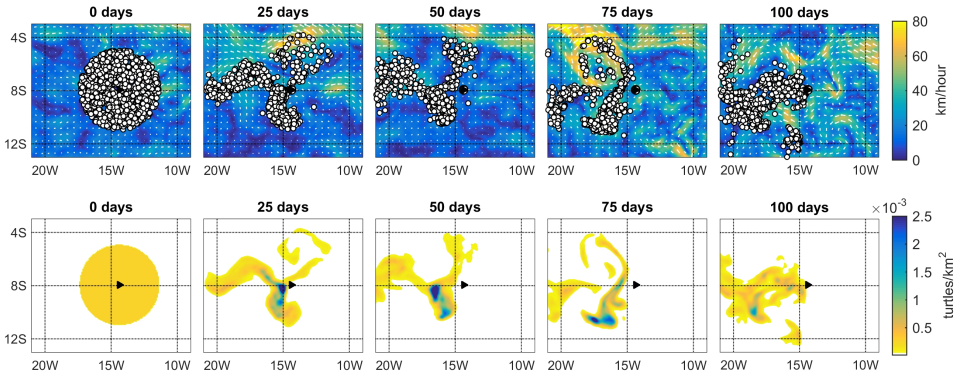


FIGURE 5. Comparison between simulations of the stochastic velocity-jump model and the fundamental solution of (6) in three dimensions. All plots are generated at time $t = 50$. In (a)-(b) we plot the three dimensional contour surface at the fixed density $c(\mathbf{x}, t) = 10^{-5}$ for (a) the fundamental solution given by equation (9) and (b) the particle probability distribution generated from multiple simulations (200000) of the 3D velocity-jump model. In (c-d) we illustrate the quantitative match by plotting contour lines of fixed density $c(\mathbf{x}, t) = 10^{-4}$ (solid line) and $c(\mathbf{x}, t) = 10^{-6}$ (dashed line) for projections onto the (c) $x - z$ plane (fixing $y = 50$) and (d) $y - z$ plane (fixing $x = 460$). For the simulations we initialise the individual at $(0, 0, 0)$ and compute the random walk path using the Fisher distribution (2) for q (with $\mathbf{u} = (1, 0, 0)$) and $\kappa = 10$) along with $s = 10$ and $\lambda = 100$. $\mathbf{a}$ and D for equation (9) are subsequently determined by equations (25) and (26).

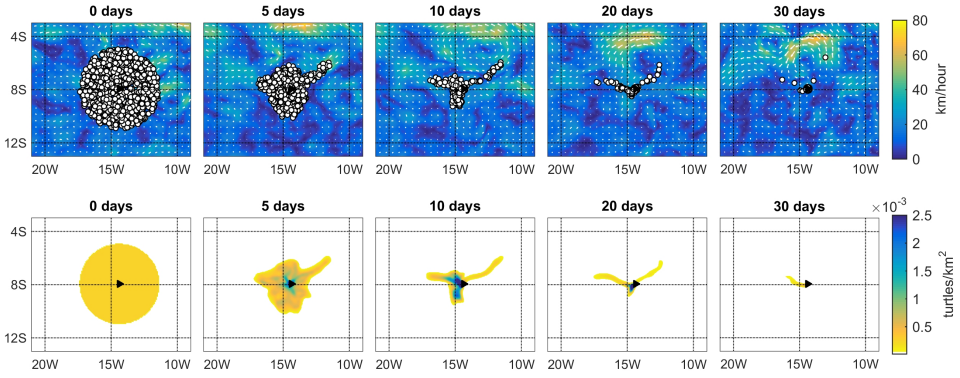
oil companies to allow seismic imaging of potential oil reservoirs. Approximately 5 metres wide, these lines criss-cross the habitats of caribou and wolf populations and have been found to impact on their spatial distribution, thereby potentially altering the predator-prey dynamics. Location data from radio collars attached to wolves and caribou found that (on average) the former tended to be closer to seismic lines and the latter further away [18], suggesting that wolves may utilise the seismic lines to facilitate their movement through the landscape (and hence expand their hunting range) while caribou avoid them, perhaps due to a greater chance of encounters with wolves. More recently, GPS collar data from wolves demonstrated that they tended to move along the seismic lines, and that a bimodal von Mises distribution provided a good fit for their movement properties [23]. Subsequent modelling in [23], based on various modes of biased movement in response to seismic lines, was used to predict how this would impact on encounter/predation rates.

To show how biased movement along seismic lines can alter the population distribution, we consider simulations of the macroscopic model (6) in different landscapes, where the expectation and variance-covariance required for $\mathbf{a}$ and D are those of the bimodal von Mises distribution (23–24). Specifically, we consider (a) an artificial landscape of regular criss-crossing seismic lines and (b) a genuine landscape based on an aerial photograph. We assume distinct movement properties on and off the lines as follows.

- Off-seismic lines the population is assumed to move in a (more or less) random manner, i.e. $k = 0$ in (4).



(a)



(b)

FIGURE 6. Simulations of green sea turtle navigation towards Ascension island: we refer to [31] for full details. We consider (a) weak navigators and (b) strong navigators, showing in each case (top row) the individual-based model and (bottom row) the macroscopic model. For the IBM turtle positions (white circles) are shown at the indicated times, superimposed on ocean currents as illustrated by its direction (arrows) and magnitude (arrow length/density map), while for the macroscopic model (6) the population density is indicated by the density map, where the scale indicates the number/km^2 and white regions represent a density below 10^{-5}. Simulations start on 1st of January 2014 and turtles swim with mean fixed swimming speed of 40 km/day and make turns once per hour. Ocean currents are obtained from a standard ocean forecasting model (HYCOM, [5]).

- On-seismic lines the population is biased into either following or avoiding the seismic lines:
  - for seismic line following we take $k = 2$ in (4) and a direction of movement along the (local) axial direction of the seismic line ;
  - for seismic line avoidance we take $k = 2$ in (4) and a dominant direction of movement orthogonal to the (local) axial direction of the seismic line.

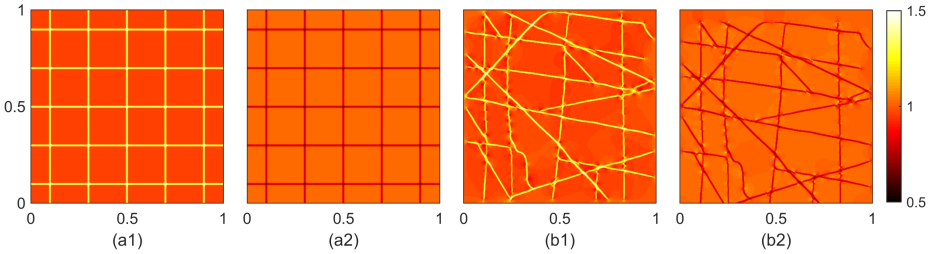


FIGURE 7. Population density distributions describing the different responses of a population to linear landscape features such as seismic lines. In each case we solve (6), where **a** and D are calculated using the moments of the bimodal von Mises distribution, i.e. equations (23–24). In (a) we impose a regular array of crisscrossing seismic lines: the population density distribution is either (a1) raised or (a2) decreased along the seismic lines according to whether they tend to (a1) follow along the paths or (a2) move directly off of them. In (b) an identical set of simulations is conducted for a genuine landscape, based on an aerial photograph of a Northern Alberta landscape.

Figure 7 plots the long-term distributions for populations following one of these scenarios. As to be expected, there is a heterogeneous pattern in the population distribution according to the underlying network of seismic lines, such that an individual is either more or less likely to be found on (or close to) a seismic line for following or avoidance strategies respectively. A more formal matching of this distribution data to the location data found in studies such as [18, 23] would allow us to examine theoretically how different strategies and responses of both wolves and caribou to the seismic line network could impact on predation rates: we refer to [23] to a specific such study in this regard.

5.4. **An application to brain tumour modelling.** One application of the three-dimensional von Mises-Fisher distribution can be found in the modelling of brain tumours. Gliomas are brain tumours that develop from the glial cells of the central nervous system. The most aggressive variant is the Grade IV Astrocytoma (Glioblastoma multiforme). Current standard of care includes a combination of maximal safe surgical de-bulking, radiation therapy, and chemotherapy. The typical median survival is approximately 14 months from diagnosis. Despite intensive treatment, recurrence is expected in virtually all patients, and usually occurs within and adjacent of the original region of grossly visible disease seen on diagnostic imaging. Microscopic extension of tumor cells can be found in the adjacent normal appearing brain and are thought to contribute to recurrence and subsequent extension of the tumour.

Brain tissue can be classified into two main components: gray matter and white matter. While the former is relatively isotropic in structure, a significant proportion of white matter stems from the myelinated nerve axons, bundled into long fibre tracts that generate a highly anisotropic tissue. Significant evidence suggests that invasive glioma cells spread along these white matter tracts [33, 11, 12] hence directing the heterogeneous spread of cancer. Consequently it is of key interest to

understand how these fibres are connected in the brain. An MRI-based technique called *diffusion tensor imaging* (DTI) generates voxel-by-voxel diffusion tensors for the anisotropic diffusion of water molecules, which can subsequently be reinterpreted to provide a measurement of tissue alignment and hence information on the arrangement of white matter tracts [4].

A number of groups have developed methods to include DTI measurements within mathematical glioma growth models [35, 13, 19, 21, 26, 30, 9] and in [30] we proposed to incorporate DTI information via its parametrisation of a bimodal von Mises-Fisher distribution. Consequently the macroscopic glioma growth follows an anisotropic diffusion model, augmented by a logistic growth term to describe proliferation of the cancer cells:

$$c(\mathbf{x},t)_t = \nabla\nabla : (D(\mathbf{x})c(\mathbf{x},t)) + rc(\mathbf{x},t)(1-c(\mathbf{x},t))$$

where $c(\mathbf{x},t)$ is the time and space dependent (normalised) cell density and r is the growth rate. The colon symbol (:) is used as contraction as in (6). $D(\mathbf{x})$ is given from (8) using (32), where the dominant direction $\mathbf{u}$ and concentration parameter k is determined by the anisotropy revealed in the DTI data and will be spatially variant. The concentration parameter k is taken to be proportional to the *fractional anisotropy* (FA) at each location, with constant of proportionality κ, i.e. $k(\mathbf{x}) = \kappa \mathrm{FA}(\mathbf{x})$. The fractional anisotropy, or FA measures "how anisotropic" the DTI measured tensor is, with FA=0 corresponding to an isotropic tensor and FA=1 corresponding to a fully anisotropic tensor. The constant κ measures the degree to which the invasion process follows the anisotropic structure and can be tuned for a specific patient: choosing $\kappa = 0$ will generate a uniform distribution, and hence an isotropic tensor, while a higher value of κ simulates a stronger tendency of the cancer cells to invade anisotropically along white matter tracts.

A full description of the model and simulation procedure can be found in [34] (see also [30]) and a typical simulation is shown in Figure 8. In (a), we plot the FA map in a two-dimensional slice of brain for a particular patient, and augment it with a black dot representing the initial condition (i.e. proposed point of tumour initiation) for a simulation; note that yellow regions correspond to high FA, and hence imply highly aligned tissue in the form of bundled white matter tracts. (b) and (d) show two-dimensional slices through the three-dimensional cell density generated by model simulation, creating an artificial tumour based on patient-specific DTI data. While the (b) and (c) corresponds to a low κ value ($\kappa = 0.5$), (d) and (e) use a high κ value ($\kappa = 15$). Here the blue to yellow colormap indicates increasing cell density and the tumour profile is plotted 280 days post initiation. (c) and (e) show an isosurface for the three-dimensional cell density corresponding to $c(x,t) = 0.16$, estimated to be the density observable on a T2 MRI image [36] and hence indicating the visible tumour boundary. We see that for higher values of κ the finger-like projections often seen in these types of tumours can be recreated.

6. **Conclusion.** We have shown that the divergence theorem can be a powerful tool to compute the moments of spherical distributions. In this paper we focussed on the first two moments of two and three dimensional spherical analogues of the Normal distribution: the von Mises distribution (in 2-D) and the Fisher distribution (in 3-D). The method presented avoids tedious trigonometric integrals and we expect that it can also be employed for higher order moments, for higher dimensions, or for other spherical distributions such as the Kent- or Bingham distributions [20, 22]. The key argument was the use of the divergence theorem, and this is also possible

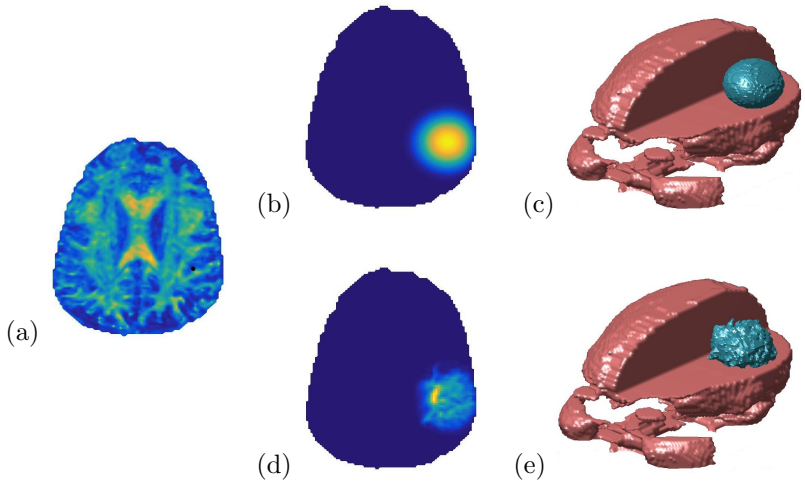


FIGURE 8. Simulations of brain tumour growth using real patient data: see [34] for details. These are test cases to show the effect of changing the concentration parameter. (a) shows the fractional anisotropy for a two-dimensional axial slice of a real patient brain. Yellow indicates high fractional anisotropy, and thus high alignment, while blue indicates isotropic tissue. The initial condition for the brain tumour simulation is indicated by a black dot. (b)-(e) show two artificial tumours generated using real patient DTI data for two different values of κ, and thus k. (b) and (c) use $\kappa = 0.5$, and (d) and (e) use $\kappa = 15$. (b) and (d) represent a two-dimensional slice of the cell density (where dark blue = low cancer cell density, yellow = high cancer cell density), while (c) and (e) represent an isosurface corresponding to $c(x,t) = 0.16$. It has been estimated that this is the cell-denstity that shows up on T2-MRI images, thus the isosurfaces in (c) and (e) can be thought of as the visible tumour boundaries.

in higher dimensions. In those cases, higher spherical moment calculations are necessary, but there is hope that these can be solved. See for example [17], where general spherical moments were computed in arbitrary space dimensions for the uniform distribution.

Our focus on the first two moments of the von Mises and Fisher distributions stems from their direct relevance to the modelling of oriented movement of animals and cells, highlighted here by applications to the modelling of sea turtle navigation, wolf movement and brain tumour invasion: such movement-based problems naturally reside in two or three dimensional space and one could apply similar methods to model a myriad of other such phenomena in biology, ecology or medicine. More generally, the n-dimensional von Mises-Fisher distribution has found applications in numerous other areas, such as describing the type of high dimensional data found in areas such as gene expression data [2]. Consequently, we hope that the method described here may prove to have wider utility for other applications requiring descriptions of directional data.

Acknowledgments. AS is grateful to support through a NSERC CGS D3 scholarship and an Alberta Innovates Technology Futures Graduate Student Scholarship. TH is supported through an NSERC discovery grant. The DTI images were acquired through an Alberta Cancer Foundation sponsored grant aimed at using DTI imaging to help predict the pattern of glioma growth.

REFERENCES

[1] wikipedia. com, https://en.wikipedia.org/wiki/Von_Mises%E2%80%93Fisher_distribution, last accessed on 11/22/2016.
[2] A. Banerjee, I. S. Dhillon, J. Ghosh and S. Sra, Clustering on the unit hypersphere using von Mises-Fisher distributions, In: *Journal of Machine Learning Research,* **6** (2005), 1345–1382.
[3] E. Batschelet, *Circular Statistics in Biology,* Academic Press, London, 1981.
[4] C. Beaulieu, The basis of anisotropic water diffusion in the nervous system - a technical review, *NMR Biomed.,* **15** (2002), 435–455.
[5] R. Bleck, An oceanic general circulation model framed in hybrid isopycnic-cartesian coordinates, *Ocean Mod.,* **4** (2002), 55–88.
[6] E. A. Codling, R. N. Bearon and G. J. Thorn, Diffusion about the mean drift location in a biased random walk, *Ecology,* **91** (2010), 3106–3113
[7] E. A. Codling, M. J. Plank and S. Benhamou, Random walk models in biology, *J. Roy. Soc. Interface,* **5** (2008), 813–834.
[8] C. Darwin, Perception in the lower animals, *Nature,* **7** (1873), 360.
[9] C. Engwer, T. Hillen, M. Knappitsch and C. Surulescu, Glioma follow white matter tracts: A multiscale DTI-based model, *J. Math. Biol.*, **71** (2015), 551–582.
[10] N. I. Fisher, *Statistical Analysis of Circular Data*, Cambridge University Press, Cambridge, 1993.
[11] A. Giese, L. Kluwe, B. Laube, H. Meissner, M. E. Berens and M. Westphal, Migration of human glioma cells on myelin, *Neurosurgery,* **38** (1996), 755–764.
[12] P. G. Gritsenko, O. Ilina and Friedl, P.: Interstitial guidance of cancer invasion, *Journal or Pathology*, **226** (2012), 185–199.
[13] H. Hatzikirou, A. Deutsch, C. Schaller, M. Simaon and K. Swanson, Mathematical modelling of glioblastoma tumour development: A review, *Math. Models Meth. Appl. Sci.*, **15** (2005), 1779–1794.
[14] T. Hillen, M^5 mesoscopic and macroscopic models for mesenchymal motion, *J. Math. Biol.*, **53** (2006), 585–616.
[15] T. Hillen, E. Leonard and H. van Roessel, *Partial Differential Equations; Theory and Completely Solved Problems,* Wiley, Hoboken, NJ, 2012.
[16] T. Hillen and K. Painter, Transport and anisotropic diffusion models for movement in oriented habitats, In: *Lewis,M., Maini,P., Petrovskii,S. (Eds.), Dispersal, Individual Movement and Spatial Ecology: A Mathematical Perspective,* Springer,Heidelberg, **2071** (2013), 177–222.
[17] T. Hillen, On the L^2-moment closure of transport equations: The general case, *Discr. Cont. Dyn. Systems, Series B*, **5** (2005), 299–318.
[18] A. R. C. James and A. K. Stuart-Smith, Distribution of caribou and wolves in relation to linear corridors, *The Journal of Wildlife Management,* **64** (2000), 154–159.
[19] A. Jbabdi, E. Mandonnet, H. Duffau, L. Capelle, K. Swanson, M. Pelegrini-Issac, R. Guillevin and H. Benali, Simulation of anisotropic growth of low-grade gliomas using diffusion tensor imaging, *Mang. Res. Med.,* **54** (2005), 616–624.
[20] J. Kent, The Fisher-Bingham Distribution on the Sphere, J. Royal. Stat. Soc., 1982.
[21] E. Konukoglu, O. Clatz, P. Bondiau, H. Delignette and N. Ayache, Extrapolation glioma invasion margin in brain magnetic resonance images: Suggesting new irradiation margins, *Medical Image Analysis,* **14** (2010), 111–125.
[22] K. V. Mardia and P. E. Jupp, *Directional Statistics,* Wiley, New York, 2000.
[23] H. McKenzie, E. Merrill, R. Spiteri and M. Lewis, How linear features alter predator movement and the functional response, *Royal Society Interface,* **2** (2012), 205–216.
[24] P. Moorcroft and M. A. Lewis, Mechanistic Home Range Analysis, Princeton University Press, USA, 2006.
[25] J. A. Mortimer and A. Carr, Reproduction and migrations of the Ascension Island green turtle (chelonia mydas), *Copeia,* **1987** (1987), 103–113.

[26] P. Mosayebi, D. Cobzas, A. Murtha and M. Jagersand, Tumor invasion margin on the Riemannian space of brain fibers, *Medical Image Analysis,* **16** (2011), 361–373.
[27] A. Okubo and S. Levin, *Diffusion and Ecological Problems: Modern Perspectives*, Second edition. Interdisciplinary Applied Mathematics, 14. Springer-Verlag, New York, 2001.
[28] H. Othmer, S. Dunbar and W. Alt, Models of dispersal in biological systems, *Journal of Mathematical Biology*, **26** (1988), 263–298.
[29] K. Painter, Modelling migration strategies in the extracellular matrix, *J. Math. Biol.*, **58** (2009), 511–543.
[30] K. Painter and T. Hillen, 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** (2013), 25–39.
[31] K. Painter and T. Hillen, Navigating the flow: Individual and continuum models for homing in flowing environments, *Royal Society Interface,* **12** (2015), 20150, 647.
[32] K. J. Painter, Multiscale models for movement in oriented environments and their application to hilltopping in butterflies, *Theor. Ecol.,* **7** (2014), 53–75.
[33] J. Rao, Molecular mechanisms of glioma invasiveness: The role of proteases, *Nature Reviews Cancer,* **3** (2003), 489–501.
[34] A. Swan, T. Hillen, J. Bowman and A. Murtha, A patient-specific anisotropic diffusion model for brain tumor spread, *Bull. Math. Biol.,* to appear (2017).
[35] K. Swanson, E. A. Jr. and J. Murray, A quantitative model for differential motility of gliomas in grey and white matter, *Cell Proliferation,* **33** (2000), 317–329.
[36] K. Swanson, R. Rostomily and E. Alvord Jr., Predicting survival of patients with glioblastoma by combining a mathematical model and pre-operative MR imaging characteristics: A proof of principle, *British Journal of Cancer,* **98** (2008), 113–119.
[37] P. Turchin, *Quantitative Analysis of Movement,* Sinauer Assoc., Sunderland, 1998.
[38] C. Xue and H. Othmer, Multiscale models of taxis-driven patterning in bacterial populations, *SIAM Journal for Applied Mathematics*, **70** (2009), 133–167.

Received May 09, 2016; Accepted November 22, 2016.

E-mail address: thillen@ualberta.ca
E-mail address: K.Painter@hw.ac.uk
E-mail address: acswan@ualberta.ca
E-mail address: albert.murtha@ahs.ca