

NEURAL NETWORKS FOR LEARNING MACROSCOPIC CHEMOTACTIC SENSITIVITY FROM MICROSCOPIC MODELS

RADEK ERBAN*

Abstract. The macroscopic (population-level) dynamics of chemotactic cell movement – arising from underlying microscopic (individual-based) models – are often described by parabolic partial differential equations (PDEs) governing the spatio-temporal evolution of cell concentrations. In certain cases, these macroscopic PDEs can be analytically derived from microscopic models, thereby elucidating the dependence of PDE coefficients on the parameters of the underlying individual-based dynamics. However, such analytical derivations are not always feasible, particularly for more complex or nonlinear microscopic models. In these instances, neural networks offer a promising alternative for estimating the coefficients of macroscopic PDEs directly from data generated by microscopic simulations. In this work, three microscopic models of chemotaxis are investigated. The macroscopic chemotaxis sensitivity is estimated using neural networks, thereby bridging the gap between individual-level behaviours and population-level descriptions. The results are compared with macroscopic PDEs, which can be derived for each model in certain parameter regimes.

Relevance to Life Sciences. Many unicellular organisms exhibit chemotaxis, moving in response to extracellular chemical cues, such as nutrients or toxins. They detect these cues via membrane-bound receptors, and binding to these receptors triggers intracellular signalling cascades that bias their movement. Stochastic, individual-based modelling can be used to capture the fine-grained dynamics at the single-cell level. In this work, we employ detailed, individual-based modelling of chemotaxis to infer population-level properties that characterise collections of cells.

Mathematical Content. At the individual level, the mathematical frameworks for chemotaxis include (i) position-jump processes (discretised stochastic differential equations), (ii) velocity-jump processes, and (iii) velocity-jump processes with internal dynamics. The first describes a cell by a stochastic process for its position, the second describes a cell by both its position and its velocity, while the third also introduces variables that describe intracellular processes. Population-level behaviour is inferred either by deriving macroscopic PDEs or by training feedforward neural networks on data obtained from simulations of microscopic, individual-based models.

Key words. chemotaxis, macroscopic equations, individual-based models, neural networks

MSC codes. 35Q92,92B20,82C31,92D25

1. Introduction. The (classical, Keller-Segel) chemotaxis equation is a partial differential equation (PDE) describing the time evolution of the spatio-temporal concentration of cells $c(\mathbf{x}, t)$ in response to a spatio-temporal chemical signal $S(\mathbf{x}, t)$. Assuming that the signal $S(\mathbf{x}, t)$ is a chemoattractant, the chemotaxis equation can be written as a drift-diffusion equation in domain $\Omega \subset \mathbb{R}^d$, for dimension $d \in \{1, 2, 3\}$, in the following form [30, 12]

$$(1.1) \quad \frac{\partial c}{\partial t} = \nabla \cdot (D \nabla c - c \chi(S) \nabla S),$$

where $c : \Omega \times [0, \infty) \rightarrow [0, \infty)$ is the concentration of cells, D is the diffusion constant and $\chi : [0, \infty) \rightarrow [0, \infty)$ is the chemotactic sensitivity.

In some applications, cells not only detect the chemical $S(\mathbf{x}, t)$ but also have the capacity to produce it as a signalling molecule or to consume it as a source of nutrition [1, 15]. Then the classical chemotaxis equation is coupled with a PDE describing the dynamics of the chemoattractant [27, 6, 35]. An example includes the original Keller-Segel system [32], which describes the chemoattractant by a parabolic PDE

*Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, United Kingdom (radek.erban@maths.ox.ac.uk).

and uses singular chemotactic sensitivity $\chi(S) = \mathcal{O}(1/S)$ as $S \rightarrow 0_+$. Such singularities are necessary to obtain travelling wave solutions in the PDE systems of the Keller-Segel type [31], which are also sometimes called the Patlak-Keller-Segel PDEs in the literature acknowledging the earlier work of Patlak [42]. Both parabolic and elliptic equations have been used for describing the chemoattractant $S(\mathbf{x}, t)$, resulting in parabolic-parabolic and parabolic-elliptic Keller-Segel PDE systems [47, 20, 3, 48]. While the chemotactic sensitivity $\chi(S)$ in some theoretical studies is assumed to be constant (independent of S), this can result in finite-time blow-up of solutions [47]. This can be prevented and the resulting PDE systems can better model biological phenomena by assuming a more realistic functional dependence of the chemotactic sensitivity $\chi(S)$ on the signal S [2, 21, 19]. For example, common assumptions include that cells are not sensitive to large signal concentrations, and it is required that $\chi(S) \rightarrow 0$ as $S \rightarrow \infty$ to get globally well-posed systems of PDEs in some models [2].

Chemotaxis can also be modelled using microscopic (individual-based) models. In this paper, we will investigate three such models, denoted as Microscopic Models I, II and III, each describing a collection of N individuals (unicellular organisms) and their responses to extracellular chemical signal S . Microscopic Model I is a signal-dependent position-jump process (Brownian dynamics), where the position of each individual is described by a stochastic differential equation (SDE). Microscopic Model II is a signal-dependent velocity-jump process, while the most complex Microscopic Model III also includes internal (intracellular) dynamics describing signal processing by each individual. We investigate the problem of inferring macroscopic PDEs, including the functional form of the chemotactic sensitivity $\chi(S)$, for each microscopic model. Such PDEs can be analytically derived under some assumptions for simplified individual-based models [39, 11, 49, 17, 37]. We will show that the macroscopic chemotaxis sensitivity can also be estimated using feedforward neural networks [23].

The paper is organised as follows. In Section 2, we introduce the chemotaxis equation, its numerical discretisation and our feedforward neural network architecture. Microscopic Models I, II and III are presented, and their macroscopic description inferred in Sections 3, 4 and 5, respectively. For each model, we will discuss parameter regimes where their macroscopic behaviour is well approximated by the parabolic chemotaxis equation (1.1). Considering velocity-jump processes, their macroscopic description can also be obtained in the form of the hyperbolic chemotaxis equation [25, 14, 9]. We will show in Sections 4 and 5 that the loss function based on the hyperbolic chemotaxis equation is more suitable for estimating the chemotactic sensitivity from relatively short simulations of individual-based models. We conclude with a discussion of our results in Section 6.

2. Chemotaxis equation and neural networks. In this paper, we will restrict our consideration to spatially one-dimensional chemotaxis models, with the macroscopic concentration of cells denoted by $c(x, t)$, where $x \in [0, L]$ for $L > 0$. In particular, domain Ω is the one-dimensional interval, $\Omega = [0, L]$, and the macroscopic chemotaxis equation (1.1) can be rewritten in the form

$$(2.1) \quad \frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} - \frac{\partial}{\partial x} \left(c \chi(S) \frac{\partial S}{\partial x} \right).$$

We will assume that the signal profile varies in space but remains independent of time, i.e. $S \equiv S(x)$. Our objective is to infer the appropriate form of chemotactic sensitivity $\chi(S)$, or, more generally, the entire macroscopic chemotaxis equation, based

on the estimation of cell concentration $c(x, t)$ by microscopic models in Sections 3, 4 and 5. We discretise interval $[0, L]$ using $n \in \mathbb{N}$ meshpoints at locations

$$(2.2) \quad \mathbf{x}_i = (i - 1/2) \Delta x, \quad i = 1, 2, \dots, n, \quad \text{where} \quad \Delta x = \frac{L}{n},$$

and denote

$$(2.3) \quad c_i(t) = c(x_i, t), \quad S_i = S(x_i).$$

Then concentrations $c_i(t)$, $i = 1, 2, \dots, n$, can be estimated in our simulations of individual-based models in Sections 3, 4 and 5 by calculating the number of individuals in interval $[(i - 1)\Delta x, i\Delta x]$ at time t . The notation (2.3) can also be used to describe a numerical method for solving the macroscopic chemotaxis equation (2.1). Using finite differences to discretise the right-hand side of equation (2.1), we obtain

$$(2.4) \quad \frac{dc_i}{dt} = \mathcal{A}_i(c, S)$$

where $\mathcal{A}_i(c, S)$ is defined by

$$(2.5) \quad \begin{aligned} \mathcal{A}_i(c, S) = & \frac{D}{\Delta x^2} (c_{i+1} - 2c_i + c_{i-1}) + \chi\left(\frac{S_{i+1} + S_i}{2}\right) \frac{(c_{i+1} + c_i)(S_i - S_{i+1})}{2\Delta x^2} \\ & + \chi\left(\frac{S_{i-1} + S_i}{2}\right) \frac{(c_{i-1} + c_i)(S_i - S_{i-1})}{2\Delta x^2}, \end{aligned}$$

for internal mesh points $i = 2, 3, \dots, n - 1$. To extend the definition of $\mathcal{A}_i(c, S)$ to boundary mesh points at $i = 1$ and $i = n$, we need to specify the boundary conditions. In our simulations of individual-based models in Sections 3, 4 and 5, we will use reflective (zero-flux) boundary conditions, which preserve the number of simulated individuals and correspond, at the macroscopic level, to boundary terms

$$(2.6) \quad \mathcal{A}_1(c, S) = \frac{D}{\Delta x^2} (c_2 - c_1) + \chi\left(\frac{S_2 + S_1}{2}\right) \frac{(c_2 + c_1)(S_1 - S_2)}{2\Delta x^2},$$

$$(2.7) \quad \mathcal{A}_n(c, S) = \frac{D}{\Delta x^2} (c_{n-1} - c_n) + \chi\left(\frac{S_{n-1} + S_n}{2}\right) \frac{(c_{n-1} + c_n)(S_n - S_{n-1})}{2\Delta x^2}.$$

Denoting the integral of χ by ψ , we have $\psi' = \chi$ and equation (2.1) can be rewritten as

$$\frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} - \frac{\partial}{\partial x} \left(c \frac{\partial \psi(S)}{\partial x} \right).$$

The steady state solution of the macroscopic chemotaxis equation (2.1) can then be written as

$$(2.8) \quad c_s(x) = \lim_{t \rightarrow \infty} c(x, t) = A \exp\left[\frac{\psi(S)}{D}\right]$$

where $A > 0$ is a constant. In particular, if the macroscopic equation is described by the chemotaxis equation (2.1), then the knowledge of the equilibrium (steady state) concentration profile $c_s(x)$ and the signal profile $S(x)$ could be used to infer the chemotactic sensitivity $\chi(S)$ by plotting the signal S against

$$\chi(S) = \frac{D c'_s(x)}{c_s(x) S'(x)},$$

where primes denote derivatives. However, such an approach would require running long-time simulations of macroscopic models until the equilibrium can be sampled. We will instead adopt an approach where we will evolve microscopic models over a specific time interval. We will consider transient time-dependent data obtained by microscopic models in Sections 3, 4 and 5, where we calculate the concentration profile $c(x, t + \tau)$ at time $t + \tau$ by evolving the microscopic model over the time interval $[t, t + \tau]$, where $\tau > 0$. This will be used to estimate the time derivative in equation (2.4) as

$$\frac{dc_i}{dt} = \frac{c_i(t + \tau) - c_i(t)}{\tau}.$$

Chemotaxis sensitivity $\chi(S)$ will be approximated by feedforward neural network architecture [23], which is schematically shown in Figure 1(a). The architecture is structured to accept a single input, the value of S , which is managed by the first fully connected layer. This layer is designed to transform the input size of one into an output size of n_1 , effectively mapping the single-dimensional input into a higher-dimensional space. This is achieved by applying an affine transformation $\mathbf{w}_1 S + \mathbf{b}_1$, where $\mathbf{w}_1 \in \mathbb{R}^{n_1}$ and $\mathbf{b}_1 \in \mathbb{R}^{n_1}$ are the weight vector and the bias vector, respectively. Following this transformation, the non-linearity is introduced through the application of a ReLU activation function defined by

$$(2.9) \quad \text{ReLU}(z) = \max\{0, z\} = \begin{cases} z & \text{for } z > 0; \\ 0 & \text{for } z \leq 0. \end{cases}$$

This non-linear activation enables the neural network to capture more complex relationships within the data.

Our neural network architecture in Figure 1(a) continues with a second hidden layer, which mirrors the structure of the first hidden layer. It takes the output from the previous layer, $\mathbf{z}_1 \in \mathbb{R}^{n_1}$, and again applies an affine transformation, which can be written as $W_2 \mathbf{z}_1 + \mathbf{b}_2$, where $W_2 \in \mathbb{R}^{n_2 \times n_1}$ is the weight matrix, $\mathbf{b}_2 \in \mathbb{R}^{n_2}$ is the bias vector, and n_2 is the size of the second hidden layer. As with the first layer, the ReLU activation function (2.9) is applied to the output of the second layer, ensuring that the model retains the ability to learn intricate, non-linear relationships. Finally, the architecture in Figure 1(a) culminates in the output layer, which is responsible for condensing the information down to a single output value. This layer takes the n_2 -dimensional input from the second hidden layer, $\mathbf{z}_2 \in \mathbb{R}^{n_2}$, and applies an affine transformation that results in a single output size of one, namely we use $\mathbf{w}_3 \cdot \mathbf{z}_2 + b_3$, where $\mathbf{w}_3 \in \mathbb{R}^{n_2}$ and $b_3 \in \mathbb{R}$ are the weight vector and the bias, respectively. To train the network, we use equation (2.4) to define the loss function

$$(2.10) \quad \mathcal{L}(S) = \frac{1}{n} \sum_{i=1}^n \left(\frac{c_i(t + \tau) - c_i(t)}{\tau} - \mathcal{A}_i(c, S) \right)^2,$$

and we use the Adam optimizer to minimize prediction errors based on the training dataset for each microscopic model. We implement our feedforward neural network architecture using PyTorch.

In our illustrative computational examples in Sections 3, 4 and 5, we will assume that all parameters have already been non-dimensionalized. We use the domain size $L = 1$ and the chemotactic signal profile $S : \Omega \rightarrow [0, 1]$ given by

$$(2.11) \quad S(x) = 2^4 x^2 (x - 1)^2.$$

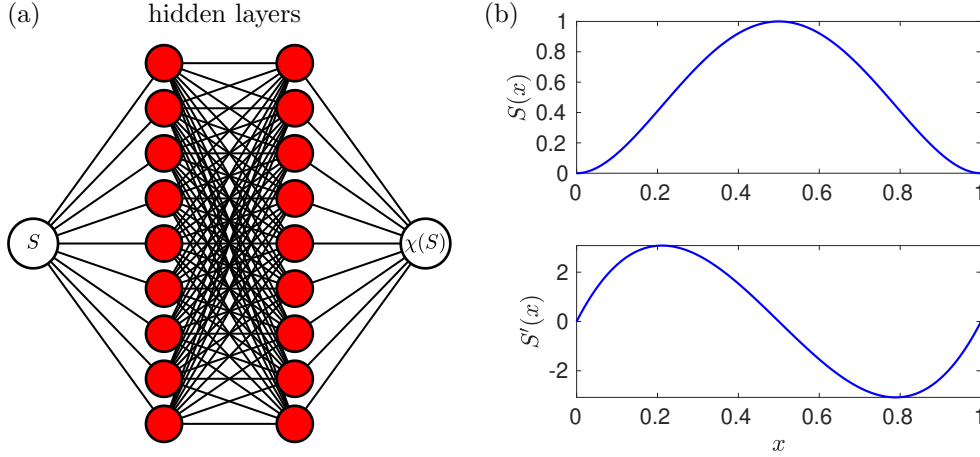


FIG. 1. (a) Schematic of the feedforward neural network architecture with two hidden layers. (b) The chemotactic signal $S(x)$ (top panel) and its derivative $S'(x)$ (bottom panel) given by equations (2.11) and (2.12).

This signal profile is plotted in the top panel of Figure 1(b). Differentiating (2.11), we obtain

$$(2.12) \quad S'(x) = 2^5 x(2x-1)(x-1) \quad \text{and} \quad S''(x) = 2^5(6x^2 - 6x + 1).$$

In particular, considering the domain $\Omega = [0, 1]$, the signal $S(x)$ has minima at $x = 0$ and $x = 1$ equal to $S(0) = S(1) = 0$ and its maximum at $x = 1/2$ equal to $S(1/2) = 1$. The derivative $S'(x)$ is plotted in the bottom panel of Figure 1(b). It satisfies

$$|S'(x)| \leq \frac{16}{3\sqrt{3}} \doteq 3.0792\dots \quad \text{for } x \in \Omega = [0, 1].$$

3. Microscopic Model I (Brownian dynamics). We model a system of N individuals (unicellular organisms) in domain $\Omega = [0, L]$, where $L > 0$. The state of the system at time t is described by the N -dimensional vector of positions

$$(3.1) \quad \mathbf{X}(t) = [X_1(t), X_2(t), \dots, X_N(t)] \in \Omega^N,$$

where $X_j(t)$ denotes the position of the j -th cell at time t for $j = 1, 2, \dots, N$. The time evolution of $X_j(t)$ is given by the following Itô SDE

$$(3.2) \quad dX_j(t) = \chi(S) \frac{\partial S}{\partial x} dt + \sqrt{2D} dW_j, \quad \text{for } j = 1, 2, \dots, N.$$

Then the chemotaxis equation (2.1) is exactly equal to the Fokker-Planck equation describing the time evolution of the spatio-temporal probability density of each individual [8]. In particular, if we simulate the microscopic model and calculate the number of individuals in interval $[(i-1)\Delta x, i\Delta x]$ at time t to estimate the density $c_i(t) = c(x_i, t)$ given by (2.3), then we are effectively calculating (noisy) solutions of the chemotaxis equation (2.1). The error (noise) reduces to zero in the large particle limit $N \rightarrow \infty$ provided that the equation (3.2) is solved exactly. In practice, we simulate a finite number of individuals, N , and we discretise the SDE (3.2) using a finite time step Δt to obtain

$$(3.3) \quad X_j(t + \Delta t) = X_j(t) + \chi(S) \frac{\partial S}{\partial x} \Delta t + \sqrt{2D \Delta t} \xi_j, \quad \text{for } j = 1, 2, \dots, N,$$

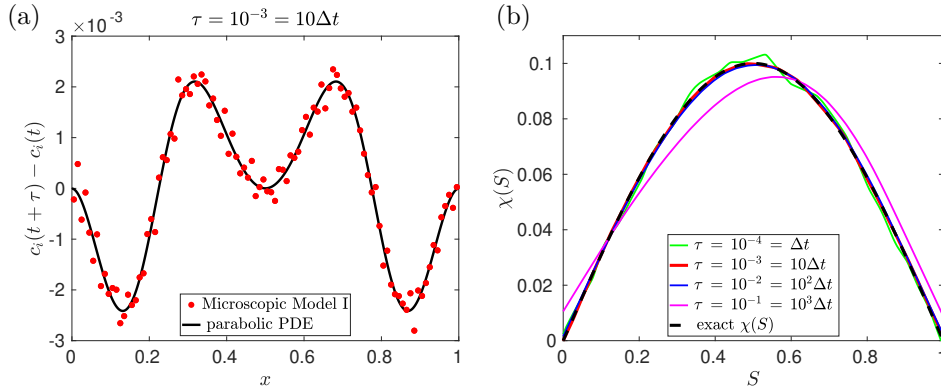


FIG. 2. (a) The change in the concentration $c_i(t+\tau) - c_i(t)$ over the time interval of length τ , starting from the uniform concentration profile $c_i(t) \equiv 1$, for $i = 1, 2, \dots, n$, calculated by Microscopic model I (red circles). We use chemotactic sensitivity (3.4), signal profile (2.11), $N = 10^9$ cells in the microscopic model, time interval $\tau = 10^{-3} = 10\Delta t$, $D = 10^{-2}$ and $n = 100$. The solution of the parabolic chemotaxis equation (2.1) is plotted as the black line.

(b) The chemotactic sensitivity estimated by the feedforward neural network for different values of time window τ . The exact chemotactic sensitivity (3.4) is plotted as the black dashed line.

where ξ_j is normally distributed random number with the zero mean and unit variance. In our simulations, we consider $N = 10^9$ individuals, $D = 10^{-2}$, $\Delta t = 10^{-4}$ and the signal profile given by (2.11) together with the chemotactic sensitivity

$$(3.4) \quad \chi(S) = \frac{\sin(\pi S)}{10}.$$

Since our signal profile (2.11) satisfies $0 \leq S(x) \leq 1$ in interval $\Omega = [0, 1]$, we have $0 \leq \chi(S) \leq 1/10$ with the highest chemotactic sensitivity achieved for intermediate signal values at $S = 1/2$. The steady state solution (2.8) is given by

$$(3.5) \quad c_s(x) = A \exp\left[\frac{-\cos(\pi S(x))}{10\pi D}\right] = A \exp\left[\frac{-\cos(2^4 \pi x^2(x-1)^2)}{10\pi D}\right],$$

where $A > 0$ is a constant. The steady state solution (3.5) has its maximum at $x = 1/2$ and concentration $c(x, t)$ will approach $c_s(x)$, provided that we observe the system for sufficiently long time. Considering transient dynamics, concentration $c(x, t)$ will, in general, not have its maximum at $x = 1/2$. This is illustrated in Figure 2(a), where we consider that cells are uniformly distributed at time t , i.e. $c_i(t) \equiv 1$, and we calculate their time evolution by (3.3) over the time interval $[t, t + \tau]$, where $\tau = 10^{-3} = 10\Delta t$. In Figure 2(a), we plot the estimated change in concentration $c_i(t + \tau) - c_i(t) \equiv c_i(t + \tau) - 1$ and compare it with the solution of the parabolic chemotaxis equation (2.1). We observe that there are initially no significant changes in concentration at points $x = 0$, $x = 1/2$ and $x = 1$, which correspond to minima (at $x = 0$ and $x = 1$) and maximum (at $x = 1/2$) of the signal $S(x)$.

Next, we use data on positions of cells calculated by relatively short simulations of Microscopic Model I to estimate the chemotactic sensitivity $\chi(S)$. We use the feedforward neural network architecture schematically shown in Figure 1(a) with $n_1 = n_2 = 50$. Our training data are calculated as in Figure 2(a), where we consider that cells are uniformly distributed at time t , i.e. $c_i(t) \equiv 1$, and we calculate their time evolution by applying equation (3.3) over the time interval $[t, t + \tau]$ to estimate $c_i(t + \tau)$,

where $\tau > 0$. We use

$$\tau \in \{\Delta t, 10\Delta t, 10^2\Delta t, 10^3\Delta t\}$$

and train 100 neural networks (using 10^3 epochs) for each pair $c_i(t)$ and $c_i(t + \tau)$. Each trained neural network provides us with one estimate of the chemotactic sensitivity $\chi(S)$. To report our results in Figure 2(b), we average these estimates, i.e. we plot $\chi(S)$ calculated as averages over 100 realizations of the neural network training process for each value of τ .

Comparing the neural network estimates of $\chi(S)$ in Figure 2(b) with the exact result (3.4), we observe that the best results are obtained when using intermediate values $\tau = 10\Delta t$ and $\tau = 10^2\Delta t$. On one hand, if τ is larger (i.e. for $\tau = 10^3\Delta t$ in Figure 2(b)), then we observe a systematic bias in Figure 2(b), caused by inaccurate estimation of the time derivative in the loss function (2.10). On the other hand, if τ is small (i.e. for $\tau = \Delta t$), then we observe a more noisy estimate of the chemotactic sensitivity in Figure 2(b), which is caused by a larger relative noise in our training data. To quantify this noise, we define the relative data error estimate

$$(3.6) \quad \varepsilon(t + \tau) = \frac{1}{n \kappa(t + \tau)} \sum_{i=1}^n \left| c_{i,\text{data}}(t + \tau) - c_i(t + \tau) \right|,$$

where $c_{i,\text{data}}(t + \tau)$ are the values of $c_i(t + \tau)$ estimated from our microscopic data, i.e. the quantities corresponding to the red circles in Figure 2(a) for $\tau = 10^{-3}$ and $N = 10^9$. Our data error estimate (3.6) is defined as the scaled ℓ^1 -norm of the difference between our microscopic simulation results, $c_{i,\text{data}}(t + \tau)$, and the solution, $c_i(t + \tau)$, of the macroscopic chemotaxis equation (2.1), which has been visualized as the black line in Figure 2(a). In Figure 3(a), we plot our data error estimate $\varepsilon(t + \tau)$, where the scaling factor $\kappa(t + \tau)$ is given by

$$(3.7) \quad \kappa(t + \tau) = \max_{i=1,2,\dots,n} c_i(t + \tau) - \min_{i=1,2,\dots,n} c_i(t + \tau).$$

This choice of scaling factor takes into account that a measure of the size of the concentration changes at time $t + \tau$ can be calculated as the difference between the maximum and minimum of the solution at time $t + \tau$, because we start with the uniformly distributed positions of cells at time t , i.e. $c_i(t) \equiv 1$. In particular, the scaling factor $\kappa(t + \tau)$, given in equation (3.7), is used to scale the data errors in equation (3.6) to calculate the relative data error estimate $\varepsilon(t + \tau)$. In Figure 3(a), we present $\varepsilon(t + \tau)$ for $\tau \in \{\Delta t, 10\Delta t, 10^2\Delta t\}$ and $N \in \{10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}\}$ in the log-log plot to illustrate that our relative data error estimate $\varepsilon(t + \tau)$ scales as $1/\sqrt{N}$ for each value of the time interval τ .

In Figure 3(a), we also see that the relative data error estimate $\varepsilon(t + \tau)$ is a decreasing function of τ for each fixed value of N . This observation can be used to explain why we have obtained less accurate results in Figure 2(b) for $\tau = \Delta t$ and $N = 10^9$. However, we can improve the accuracy of the neural network estimation of the chemotactic sensitivity $\chi(S)$ for $\tau = \Delta t$ by increasing the value of N , because this decreases the value of the relative data error $\varepsilon(t + \tau)$, as we see for $N = 10^{10}$ in Figure 3(a).

The results reported in Figure 2(b) are averages over 100 realizations of the training process for each value of τ . If we instead reported the results of a single realization of the neural network training, then we would obtain almost the same curves for $\tau \in \{10\Delta t, 10^2\Delta t, 10^3\Delta t\}$, because the small relative data error $\varepsilon(t + \tau)$ leads to a similar estimate of the chemactic sensitivity $\chi(S)$ by each trained neural network.

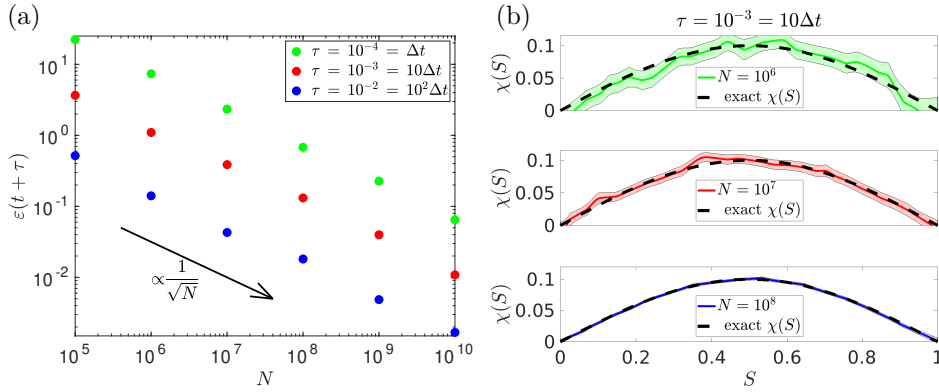


FIG. 3. (a) Relative data error estimate $\varepsilon(t + \tau)$ for time windows $\tau \in \{\Delta t, 10\Delta t, 10^2\Delta t\}$ and $N \in \{10^5, 10^6, 10^7, 10^8, 10^9, 10^{10}\}$ in the log-log plot. The black arrow illustrates the $1/\sqrt{N}$ scaling of the data error with N . We use Microscopic model I with the chemotactic sensitivity (3.4), signal profile (2.11), $D = 10^{-2}$, and $\varepsilon(t + \tau)$ given by (3.6)–(3.7) for $n = 100$. (b) The chemotactic sensitivity estimated by the feedforward neural network for three different values of $N \in \{10^6, 10^7, 10^8\}$ and time window $\tau = 10^{-3} = 10\Delta t$. The results are calculated as averages over 10^3 neural network trainings with standard deviations indicated by the shaded areas. The exact chemotactic sensitivity (3.4) is plotted as the black dashed line.

This can be further quantified by calculating the standard deviation of the estimates of the chemotactic sensitivity. For example, the exact value of $\chi(S)$ at $S = 1/2$ is equal to $\chi(1/2) = 10^{-1}$ and the standard deviation of the estimated $\chi(1/2)$ by 100 neural networks is 4.7×10^{-4} for $\tau = 10\Delta t$, 2.8×10^{-4} for $\tau = 10^2\Delta t$ and 4.2×10^{-4} for $\tau = 10^3\Delta t$. In particular, the relative error of the estimation of $\chi(1/2)$ is less than 0.5% in each of the three cases $\tau \in \{10\Delta t, 10^2\Delta t, 10^3\Delta t\}$. On the other hand, this relative error increases to 3.9%, when we consider the results calculated in Figure 2(b) for $\tau = \Delta t$. This is caused by the higher value of the relative data error $\varepsilon(t + \tau)$ for $\tau = \Delta t$.

In Figure 3(b), we plot not only the average, but also the standard deviation. In this plot, we use $\tau = 10\Delta t$, the value for which we had a negligible standard deviation for $N = 10^9$. We decrease the value of N to $N = 10^6$ (top panel), $N = 10^7$ (middle panel) and $N = 10^8$ (bottom panel in Figure 3(b)), and we illustrate that the standard deviation increases as we decrease N . However, we note that we have only used a single calculation of the microscopic dynamics in the time window $[t, t + \tau]$ to train the neural networks. In particular, if we have a lower value of N , say $N = 10^6$, then we can improve the training process by using data from multiple simulations of the microscopic dynamics in the time window $[t, t + \tau]$. In this case, we can obtain a comparably accurate results by using 10^3 realizations of $c_i(t + \tau)$ with $N = 10^6$, as we have reported in Figure 2(b) by using only one realization of $c_i(t + \tau)$ with $N = 10^9$ cells. In general, if we have a smaller number of cells, N , in our model and we do not want to modify N , then we can improve the quality of the input data and the accuracy of our neural network estimation of the chemotactic sensitivity $\chi(S)$ by considering more pairs of $c_i(t)$ and $c_i(t + \tau)$ to train the neural network.

4. Microscopic Model II (velocity-jump process). To introduce more microscopic detail into the modelling of chemotaxis, we will describe individual cells not only by their positions (3.1), but also by their velocities. Then we can view chemotaxis at the individual-based level as a velocity-jump process [39, 26]. Velocity-jump processes have been used for modelling the chemotaxis of flagellated bacteria, such as *E. coli*, which alternates between two modes of behaviour: a more or less straight motion with constant speed called “run” and a highly erratic motion called “tumble”, which produces little translation but reorients the cell [11]. Since the tumble time is shorter than the average run time, a tumble can be viewed as an instantaneous change in velocity. In this section, we again restrict to one spatial dimension and we model a system of N unicellular organisms in domain $\Omega = [0, L]$, where the state of the system at time t is described by two N -dimensional vectors: the vector of positions (3.1) and the vector of velocities denoted by

$$(4.1) \quad \mathbf{V}(t) = [V_1(t), V_2(t), \dots, V_N(t)] \in \mathbb{R}^N,$$

where $V_j(t)$ is the velocity of the j -th cell at time t for $j = 1, 2, \dots, N$. We assume that each cell moves along the x -axis at a constant speed $\beta > 0$, i.e. its velocity $V_j(t)$ only takes one of two possible values, $V_j(t) = \pm\beta$. Each cell reverses its direction at random instants of time according to the Poisson process with turning frequency

$$(4.2) \quad \text{ReLU}\left(\lambda_0 \pm \frac{\lambda_0 \chi(S)}{\beta} \frac{\partial S}{\partial x}\right),$$

where $\lambda_0 > 0$, the function ReLU is defined by (2.9) and the sign \pm depends on the direction of the cell movement: a plus sign is for the individuals moving to the left and a minus sign is for the individuals moving to the right. This choice of turning frequency ensures that the underlying velocity-jump process is biased in a way that the cell is less likely to change direction when moving in a favourable direction, i.e. in the direction in which the signal function S is increasing. Taking into account that the velocity of each individual, $V_j(t)$, takes only two possible values $V_j(t) = \pm\beta$, the turning frequency (4.2) can also be equivalently rewritten as

$$(4.3) \quad \text{ReLU}\left(\lambda_0 \left(1 - \frac{\chi(S(X_j(t)))}{V_j(t)} \frac{\partial S}{\partial x}(X_j(t))\right)\right),$$

where \pm sign has been adsorbed into the sign of the velocity $V_j(t)$ and we have also explicitly indicated that the signal and its derivative are evaluated at the current position of the individual, $X_j(t)$. The formula (4.2) includes the function ReLU to ensure that the turning frequency is nonnegative. However, if the signal $S(x)$ satisfies

$$(4.4) \quad \left| \frac{\chi(S)}{\beta} \frac{\partial S}{\partial x} \right| \leq 1,$$

then we can drop the function ReLU from the turning frequency (4.2), because its argument is always nonnegative. Moreover, let $c^+(x, t)$ be the concentration of cells that are at (x, t) and are moving to the right, and let $c^-(x, t)$ be the concentration of cells that are at (x, t) and are moving to the left. Then $c^\pm(x, t)$ satisfy the equations

$$(4.5) \quad \frac{\partial c^+}{\partial t} + s \frac{\partial c^+}{\partial x} = - \left(\lambda_0 - \frac{\lambda_0 \chi(S)}{\beta} \frac{\partial S}{\partial x} \right) c^+ + \left(\lambda_0 + \frac{\lambda_0 \chi(S)}{\beta} \frac{\partial S}{\partial x} \right) c^-,$$

$$(4.6) \quad \frac{\partial c^-}{\partial t} - s \frac{\partial c^-}{\partial x} = \left(\lambda_0 - \frac{\lambda_0 \chi(S)}{\beta} \frac{\partial S}{\partial x} \right) c^- - \left(\lambda_0 + \frac{\lambda_0 \chi(S)}{\beta} \frac{\partial S}{\partial x} \right) c^+.$$

The concentration of cells at (x, t) is given by the sum $c(x, t) = c^+(x, t) + c^-(x, t)$. Adding and subtracting the equations (4.5)–(4.6), we get

$$(4.7) \quad \frac{\partial c}{\partial t} + \beta \frac{\partial q}{\partial x} = 0,$$

$$(4.8) \quad \frac{\partial q}{\partial t} + \beta \frac{\partial c}{\partial x} = -2\lambda_0 q + \frac{2\lambda_0 \chi(S)}{\beta} \frac{\partial S}{\partial x} c,$$

where $q(x, t) = c^+(x, t) - c^-(x, t)$. Differentiating equation (4.7) with respect to t and equation (4.8) with respect to x , we can eliminate q to deduce that $c(x, t)$ satisfies the second order PDE

$$(4.9) \quad \frac{1}{2\lambda_0} \frac{\partial^2 c}{\partial t^2} + \frac{\partial c}{\partial t} = D \frac{\partial^2 c}{\partial x^2} - \frac{\partial}{\partial x} \left(c \chi(S) \frac{\partial S}{\partial x} \right),$$

where we have denoted

$$(4.10) \quad D = \frac{\beta^2}{2\lambda_0}.$$

The macroscopic equation (4.9) is a hyperbolic version of the classical chemotaxis equation (2.1) containing an additional term with the second time derivative. In particular, if we simulate Microscopic Model II and calculate the number of individuals in interval $[(i-1)\Delta x, i\Delta x]$ at time t to estimate the density $c_i(t) = c(x_i, t)$ given by (2.3), then we are effectively calculating (noisy) solutions of the hyperbolic chemotaxis equation (4.9). The error (noise) reduces to zero in the large particle limit $N \rightarrow \infty$ provided that the Poisson process with the turning frequency (4.2) is implemented exactly and the signal S satisfies (4.4). Otherwise, we will only obtain an approximation of a solution to the equation (4.9). Since the hyperbolic PDE (4.9) reduces for sufficiently large times to the parabolic chemotaxis equation, we can also use Microscopic Model II to estimate the solutions of the classical chemotaxis equation (2.1), or its steady state behaviour [28].

To solve equation (4.9) numerically, we discretise interval $[0, L]$ using $n \in \mathbb{N}$ meshpoints at locations (2.2) and use notation (2.3) as we have done when solving the parabolic chemotaxis equation (2.1) in Section 2. Using finite differences to discretise the right-hand side of equation (4.9), we obtain

$$(4.11) \quad \frac{1}{2\lambda_0} \frac{d^2 c_i}{dt^2} + \frac{dc_i}{dt} = \mathcal{A}_i(c, S),$$

where $\mathcal{A}_i(c, S)$, for $i = 1, 2, \dots, n$, is defined by (2.5)–(2.7). The system of n second-order ordinary differential equations (ODEs) (4.11) can be rewritten as a system of $2n$ first-order ODEs

$$(4.12) \quad \frac{dc_i}{dt} = z_i$$

$$(4.13) \quad \frac{dz_i}{dt} = 2\lambda_0 (\mathcal{A}_i(c, S) - z_i)$$

where the auxiliary variable z_i is the time derivative of c_i . Alternatively, we could also use other sets of two variables to rewrite the second-order PDE (4.9) as a system of two first-order PDEs, as it is shown with variables c^+ and c^- in equations (4.5)–(4.6) and variables c and q in equations (4.7)–(4.8).

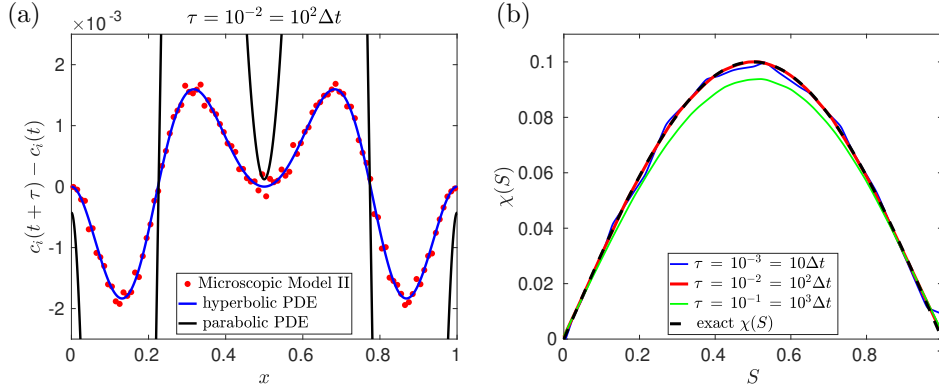


FIG. 4. (a) The change in the concentration $c_i(t+\tau) - c_i(t)$ over the time interval of length τ , starting from the uniform concentration profile $c_i(t) \equiv 1$ for $i = 1, 2, \dots, n$, calculated by Microscopic Model II (red circles). We use chemotactic sensitivity (3.4), signal profile (2.11), $N = 10^9$ cells in the microscopic model, time interval $\tau = 10^{-2} = 10^2 \Delta t$, $\beta = 0.4$, $\lambda_0 = 8$ and $n = 100$. The solutions of the parabolic chemotaxis equation (2.1) and the hyperbolic chemotaxis equation (4.9) are plotted as the black and blue line, respectively. (b) The chemotactic sensitivity estimated by the feedforward neural network for different values of the time interval τ . The exact chemotactic sensitivity (3.4) is plotted as the black dashed line.

To simulate Microscopic Model II, we update the system over time steps of length Δt . Multiplying the turning frequency (4.3) by Δt and assuming that Δt is sufficiently small, the probability that the j -th cell changes its direction during one time step is given by

$$(4.14) \quad p_j(t) = \text{ReLU} \left(\lambda_0 \left(1 - \frac{\chi(S(X_j(t)))}{V_j(t)} \frac{\partial S}{\partial x}(X_j(t)) \right) \right) \Delta t.$$

In our simulations, we generate N random numbers r_j , $j = 1, 2, \dots, N$, uniformly distributed in interval $[0, 1]$, and the j -th cell changes its direction of movement at time t if $r_j < p_j(t)$. Then we update the position $X_j(t)$ of each cell by

$$(4.15) \quad X_j(t + \Delta t) = X_j(t) + V_j(t) \Delta t,$$

where $V_j(t)$ is assumed to be constant during the time interval $[t, t + \Delta t]$, i.e. we have $V_j(t) = \pm\beta$ with the sign depending on the direction of movement.

In Figure 4(a), we present illustrative results calculated by Microscopic Model II. We consider $N = 10^9$ cells, $\beta = 0.4$, $\lambda_0 = 8$, $\Delta t = 10^{-4}$, $L = 1$ and the signal profile given by (2.11) together with the chemotactic sensitivity $\chi(S)$ given by (3.4). In particular, equation (4.10) implies that $D = 10^{-2}$, which means that we use the same diffusion constant D and chemotactic signal $S(x)$ as we used in our previous simulation of Microscopic Model I in Figure 2(a). We consider the same initial condition for positions, uniform distribution $c_i(t) \equiv 1$, and we initialize a half of the cells with the positive velocity, $+\beta$, and a half of them with the negative velocity, $-\beta$, so that we have $z_i(t) \equiv 0$ as the initial condition of equations (4.12)–(4.13) at time t . We calculate the evolution over the time interval $[t, t + \tau]$, where $\tau = 10^{-2} = 10^2 \Delta t$. We note that this value of τ is 10-times longer than the value of τ used in Figure 2(a). In particular, the solution of the parabolic chemotaxis equation (2.1) varies more in Figure 4(a) than in Figure 2(a). However, it is not the parabolic chemotaxis equation (2.1) but the hyperbolic chemotaxis equation (4.9), which approximates well

Microscopic Model II. This is illustrated in Figure 4(a), where some parts of the solution of the parabolic chemotaxis equation (2.1) are outside of the plotted range of values.

The solutions of macroscopic PDEs (2.1) and (4.9) converge as $t \rightarrow \infty$ to the same steady state solution (2.8). In particular, if τ was sufficiently large (approximately of the order $\tau = 10 = 10^5 \Delta t$ for our parameter values), then the black and blue solid lines in Figure 4(a) would be close to each other and close to the steady state solution (3.5). However, our neural network estimation uses relatively small values of τ and Figure 4(a) highlights a large difference between the transient solutions of macroscopic PDEs (2.1) and (4.9) for $\tau = 10^{-2} = 10^2 \Delta t$. Therefore, we will modify the loss function (2.10) to include the central difference approximation of the additional term describing the second time derivative in equation (4.9). We use

$$(4.16) \quad \mathcal{L}_2(S) = \frac{1}{n} \sum_{i=1}^n \left(\frac{c_i(t+2\tau) - 2c_i(t+\tau) + c_i(t)}{2\lambda_0\tau^2} + \frac{c_i(t+2\tau) - c_i(t)}{2\tau} - \mathcal{A}_i(c, S) \right)^2,$$

where the spatial derivative term $\mathcal{A}_i(c, S)$ is evaluated using the calculated data $c_i(t+\tau)$ at time $t+\tau$. To estimate the chemotactic sensitivity $\chi(S)$ from relatively short simulations of Microscopic Model II, we use the feedforward neural network architecture schematically shown in Figure 1(a) with $n_1 = n_2 = 50$. Our data are calculated as in Figure 4(a), where we consider that cells are uniformly distributed at time t , i.e. $c_i(t) \equiv 1$. We calculate their time evolution by Microscopic Model II over the time interval $[t, t+2\tau]$ to estimate $c_i(t+\tau)$ and $c_i(t+2\tau)$, where $\tau > 0$. We use $\tau \in \{10\Delta t, 10^2\Delta t, 10^3\Delta t\}$ and train 100 neural networks (using 10^3 epochs) for each triplet $c_i(t)$, $c_i(t+\tau)$ and $c_i(t+2\tau)$. The results are presented in Figure 4(b), where we average the calculated chemotactic sensitivity $\chi(S)$ over 100 realizations of the training process for each value of τ .

5. Microscopic Model III (with internal dynamics). When bacteria *E. coli* move in a favourable direction (that is, in the direction of increasing chemoattractant S) the run times are increased and the turning frequency of the velocity jump process decreases. This has been modelled in Section 4 by the turning frequency in the form (4.2), which assumes that a cell can directly estimate the gradient of the chemoattractant concentration at its current position.

In more detailed individual-based models of chemotaxis, *E. coli* does not directly estimate the gradient, but only the absolute value of the chemoattractant at its current position, $S(X_j(t))$. *E. coli* detects the attractant concentration by receptors on its membrane and the information about the receptor occupancy is then processed by intracellular signalling molecules forming the signal transduction network [5, 44]. Therefore, each bacterium is described not only by its position and velocity, but also by additional, internal, variables. Simplified models of this process were studied in [11, 12] which showed that models with one or two internal variables can be used to describe the excitation-adaptation properties of the intracellular signal transduction network. We will study such a one-dimensional model in this paper as Microscopic Model III.

We model a system of N unicellular organisms in domain $\Omega = [0, L]$, where the state of the system at time t is described by three N -dimensional vectors: the vector of positions (3.1), the vector of velocities (4.1) and the vector of internal variables denoted by

$$(5.1) \quad \mathbf{Y}(t) = [Y_1(t), Y_2(t), \dots, Y_N(t)] \in \mathbb{R}^N,$$

where $Y_j(t) \in \mathbb{R}$ is the internal variable of the j -th cell at time t for $j = 1, 2, \dots, N$, which evolves according to the ODE

$$(5.2) \quad \frac{dY_j}{dt} = \frac{\psi(S(X_j(t))) - Y_j}{t_a},$$

where $\psi' = \chi$ is the integral of the chemotactic sensitivity and $t_a > 0$ is the adaptation time of the signal transduction network [11]. We again assume that each cell moves along the x -axis at a constant speed $\beta > 0$, i.e. its velocity $V_j(t)$ only takes one of two possible values, $V_j(t) = \pm\beta$. Each cell reverses its direction at random instants of time according to the Poisson process with turning frequency

$$(5.3) \quad \text{ReLU} \left(\lambda_0 + \frac{\lambda_0 (1 + 2t_a) (Y_j - \psi(S))}{t_a \beta^2} \right),$$

where $\lambda_0 > 0$, the function ReLU is defined by (2.9) and signal S is evaluated at the current position of the cell, $X_j(t)$. Since the equation (5.2) is solved along the trajectory, the biological meaning of internal variables is that they form a primitive chemical “memory” enabling a comparison of the cell’s current environment with the environment it visited a while ago. The turning frequency (5.3) is chosen as the constant λ_0 plus the second term which depends on the internal variable Y_j . This is similar to the turning frequency (4.2). If the concentration of chemoattractant S is increasing along the cellular trajectory, then the bacterium is moving in a good direction and it is less likely to change it. On the other hand, if the concentration of S is decreasing along its trajectory, then the bacterium is more likely to turn.

To simulate Microscopic Model III, we update the system over time steps of length Δt . Multiplying the turning frequency (5.3) by Δt and assuming that Δt is sufficiently small, the probability of turning during one time step is given by

$$(5.4) \quad p_j(t) = \text{ReLU} \left(\lambda_0 + \frac{\lambda_0 (1 + 2t_a) (Y_j - \psi(S(X_j(t))))}{t_a \beta^2} \right) \Delta t.$$

In our simulations, we generate N random numbers r_j , $j = 1, 2, \dots, N$, uniformly distributed in interval $[0, 1]$, and the j -th cell changes its direction of movement at time t if $r_j < p_j(t)$. Then we update the position $X_j(t)$ of each cell by (4.15). To calculate $Y_j(t + \Delta t)$, we need to solve the ODE (5.2) over the time interval $[t, t + \Delta t]$. Since $V_j(t)$ is assumed to be constant during the time interval $[t, t + \Delta t]$, equation (5.2) reduces to solving

$$(5.5) \quad \frac{dY_j}{d\sigma} = \frac{\psi(S(X_j(t) + \sigma V_j(t))) - Y_j}{t_a}, \quad \text{for } \sigma \in [0, \Delta t],$$

given the initial condition $Y_j(t)$ for $\sigma = 0$.

Equation (5.5) can be solved by a numerical method for solving ODEs, which is often the case when using complex ODE models of signal transduction networks [5, 44]. In our case, we only have one single ODE (5.5), so we can also proceed with an analytic approach, which will highlight the hidden dependence of Microscopic Model III on the signal derivative. To see this, we observe that $\sigma \in [0, \Delta t]$ is small and we use the Taylor expansion to approximate

$$\psi(S(X_j(t) + \sigma V_j(t))) \approx \psi(S(X_j(t))) + \chi(S(X_j(t))) \frac{\partial S}{\partial x}(X_j(t)) \sigma V_j(t),$$

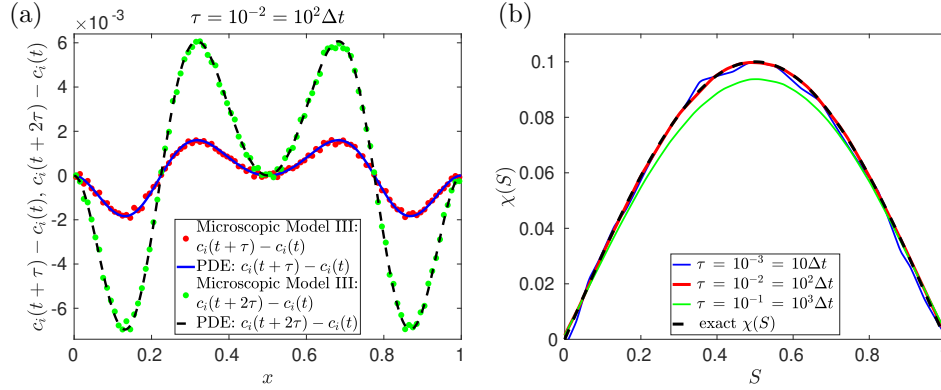


FIG. 5. (a) The changes in the concentration $c_i(t+\tau) - c_i(t)$ and $c_i(t+2\tau) - c_i(t)$ over the time intervals of length τ and 2τ , respectively, starting from the uniform concentration profile $c_i(t) \equiv 1$ for $i = 1, 2, \dots, n$, calculated by Microscopic Model III (circles). We use chemotactic sensitivity (3.4), signal profile (2.11), $N = 10^9$ cells in the microscopic model, time interval $\tau = 10^{-2} = 10^2 \Delta t$, $\beta = 0.4$, $\lambda_0 = 8$, $t_a = 10^{-4}$ and $n = 100$. The solutions of the hyperbolic chemotaxis equation (4.9) are plotted as the solid blue and dashed black lines, respectively. (b) The chemotactic sensitivity estimated by the feedforward neural network for different values of the time interval τ . The exact chemotactic sensitivity (3.4) is plotted as the black dashed line.

where $\psi' = \chi$ and we neglected terms of the order $\mathcal{O}(\sigma^2)$. Substituting into (5.5) and solving the resulting linear ODE, we obtain

$$(5.6) \quad Y_j(t + \Delta t) = Y_j(t) \exp\left(-\frac{\Delta t}{t_a}\right) + \psi(S(X_j(t))) \left(1 - \exp\left(-\frac{\Delta t}{t_a}\right)\right) + \chi(S(X_j(t))) \frac{\partial S}{\partial x}(X_j(t)) V_j(t) \left(\Delta t - t_a + t_a \exp\left(-\frac{\Delta t}{t_a}\right)\right),$$

which we use in our implementation of Microscopic Model III to calculate the update of the internal variable over one time step $[t, t + \Delta t]$. The formula (5.6) not only gives us a possible way how we can calculate $Y_j(t + \Delta t)$ at time $t + \Delta t$ given the values $X_j(t)$, $V_j(t)$ and $Y_j(t)$ at time t , but it also illustrates that cells effectively extract some information on the chemotactic sensitivity $\chi = \psi'$ and the signal derivative even if the formulation of Microscopic Model III assumes that cells can only measure the absolute values of the signal S in their environment.

Using [11, equation (6.34)], we can also derive the hyperbolic chemotaxis equation in the form (4.9), which gives a good approximation of the macroscopic behaviour of Microscopic Model III provided that the signal gradients are sufficiently shallow. This derivation is further discussed in Section 6, where we introduce a generalization of Microscopic Model III given by equation (6.1). In Figure 5(a), we present illustrative results calculated by Microscopic Model III and by solving the hyperbolic chemotaxis equation (4.9). We consider $N = 10^9$ cells, $\beta = 0.4$, $\lambda_0 = 8$, $\Delta t = 10^{-4}$, $t_a = 10^{-4}$, $L = 1$ and the signal profile given by (2.11) together with the chemotactic sensitivity $\chi(S)$ given by (3.4). In particular, equation (4.10) implies that $D = 10^{-2}$, which means that we use the same diffusion constant D , turning frequency λ_0 and chemotactic signal $S(x)$ as we used in our previous simulation of Microscopic Model II in Figure 4(a). We also consider the same initial condition for positions and velocity, namely uniform distribution $c_i(t) \equiv 1$, and we initialize a half of the cells with the positive velocity, $+\beta$, and a half of them with the negative velocity, $-\beta$, so that

we have $z_i(t) \equiv 0$ as the initial condition of equations (4.12)–(4.13) at time t . To initialize internal dynamics variables, we put $Y_j(0) = \psi(S(X_j(0)))$. We calculate the evolution over the time interval $[t, t + 2\tau]$, where $\tau = 10^{-2} = 10^2 \Delta t$ and we plot both $c_i(t + \tau) - c_i(t)$ and $c_i(t + 2\tau) - c_i(t)$ in Figure 5(a), which are the training data needed for our neural network estimation of the chemotactic sensitivity when using the loss function $\mathcal{L}_2(S)$ given by (4.16).

In Figure 5(b), we present results of the estimation of the chemotactic sensitivity $\chi(S)$ from relatively short simulations of Microscopic Model III, using the feedforward neural network architecture schematically shown in Figure 1(a) with $n_1 = n_2 = 50$. Our data are calculated as in Figure 5(a), where we consider that cells are uniformly distributed at time t , i.e. $c_i(t) \equiv 1$. We calculate their time evolution by Microscopic Model III over the time interval $[t, t + 2\tau]$ to estimate $c_i(t + \tau)$ and $c_i(t + 2\tau)$, where $\tau > 0$, which we use in the loss function (4.16). Using $\tau \in \{10\Delta t, 10^2\Delta t, 10^3\Delta t\}$, we train 100 neural networks (using 10^3 epochs) for each triplet $c_i(t)$, $c_i(t + \tau)$ and $c_i(t + 2\tau)$. The results shown in Figure 5(b) present the average calculated chemotactic sensitivity $\chi(S)$ over 100 realizations of the training process for each value of τ .

6. Discussion. Movement of individual biological agents often depends on external signal fields (for example, cells are attracted to nutrients and repelled by toxins), and population-level mathematical models of their behaviour have traditionally been formulated in terms of PDEs, such as the parabolic chemotaxis equation (1.1). At the individual level, stochastic models have been used, including Brownian dynamics and velocity-jump processes, and an important mathematical question has been to connect these two levels of description [41]. In this paper, we have investigated three individual-based models, denoted Microscopic Models I, II and III. In each case, the cellular population (in the limit of many individuals, $N \rightarrow \infty$) can be described by the macroscopic chemotaxis PDE of the form (1.1). However, it is only for the simplest Microscopic Model I that the limit $N \rightarrow \infty$ can be established analytically for all parameter regimes, and the individual-based description, formulated as Brownian dynamics with finite N and a finite time step Δt , converges to the PDE description in the limit $N \rightarrow \infty$ and $\Delta t \rightarrow 0$.

Microscopic Models II and III are formulated as velocity-jump processes in which the turning frequency depends on the extracellular signal S , either directly, as in (4.2), or indirectly through internal dynamics, as in (5.3). Then the classical chemotaxis equation can only be analytically derived under certain additional assumptions, for example, when the signal gradients, are sufficiently small [11, 49]. Microscopic Model III has included a simplified description of internal dynamics, using only one internal variable $Y_j(t)$ describing the adaptation dynamics. To model the excitation-adaptation dynamics of the intracellular signal transduction networks [40, 12], we can use a model with two variables $Y_j(t)$ and $Z_j(t)$

$$(6.1) \quad \frac{dZ_j}{dt} = \frac{g(C(t)) - Z_j - Y_j}{t_e}, \quad \frac{dY_j}{dt} = \frac{g(C(t)) - Y_j}{t_a},$$

where $C(t) = S(\mathbf{X}_j(t), t)$ is the signal detected by a cell at its current position $\mathbf{X}_j(t)$, function $g : [0, \infty) \rightarrow [0, \infty)$ incorporates sensitivity of the signal transduction network to different levels of the detected signal and constants t_e and t_a are the excitation time and the adaptation time, respectively. If $C(t)$ is a constant, then the steady state value of Z_j is independent of this constant, i.e. the values of Z_j perfectly adapt to any constant stimulus, and the velocity-jump description of chemotaxis can be formulated by assuming that the turning frequency depends on Z_j . For example, if the

turning frequency is given by $\text{ReLU}(\lambda_0 - b_0 Z_j)$ for constants $\lambda_0 > 0$ and $b_0 > 0$, then model (6.1) leads to the hyperbolic chemotaxis equation of the form (4.9), provided that the signal gradients are sufficiently small, with the chemotactic sensitivity given by [11, 12]

$$\chi(S) = \frac{b_0 \beta^2 t_a g'(S)}{\lambda_0(1 + 2\lambda_0 t_a)(1 + 2\lambda_0 t_e)}.$$

Since the adaptation time t_a is much larger than the excitation time t_e , one can also work with the simplified model, where $t_e = 0$, which has been our case of Microscopic Model III. We have studied one-dimensional models, but Microscopic Models I, II and III can be formulated in two and three spatial dimensions, and it is possible to establish analytically links with the chemotaxis equation (1.1), but the derivation is more technical [24, 12] and the simplifying assumptions of shallow signal gradients must again be imposed for Microscopic Models II and III. This analysis can be further extended to velocity-jump processes which assume that the change in the direction of motion is not instantaneous, including models of the tumbling phase of *E. coli* motion [11], and other delays caused by turning of individuals [45].

More complex chemotaxis models may include dozens of internal variables representing the concentrations of intracellular signalling molecules involved in signal transduction networks [5, 44], or written as PDEs describing spatio-temporal concentrations of signalling molecules [13, 22]. Some models also incorporate interactions between individuals, which can be direct – via volume exclusion, whereby only a finite number of cells can occupy a given volume – or indirect, by modifying the extracellular environment through the release or consumption of signalling molecules. In this case, the time evolution of extracellular signals must also be modelled. A number of individual-based models of interacting cells have been developed in the literature [7, 16, 18, 50]. The key aspect of these hybrid models is the simulation of cells as individual particles while describing the extracellular environment by PDEs. The same situation also arises when using Brownian dynamics of charged particles, where models of individual particles need to be coupled with the description of electrostatic potential solving the Poisson equation [29, 51]. Some models of cell migration also include domain growth, which introduces additional terms into population-level PDEs [4, 36, 33]. Considering complex chemotaxis models, computational frameworks for extracting macroscopic behaviour from individual-based descriptions are often necessary in some parameter regimes, where macroscopic equations cannot be analytically derived [10, 34, 43].

In this paper, we have used *in silico* data obtained by simulating Microscopic Models I, II and III to train feedforward neural networks for estimating the chemotactic sensitivity $\chi(S)$. Our training data are subject to different levels of noise, stemming from the fact that we simulate a finite number of cells, N , which are described by the same microscopic model. Other sources of noise can arise from variability between cells in their responses to the chemotactic signals [46] or by experimental error if our data are obtained through experimental chemotaxis studies [38, 52]. Our neural network approach is, in general, applicable to any of these cases, provided that our data can be used to estimate noisy concentration profiles. In Figure 3(a), we have illustrated a typical scaling of noise in our computational studies. Higher levels of data noise correspond to higher variance of the results obtained from different realizations of the neural network training process, as illustrated in Figure 3(b). Larger variance is a good indication that we do not have enough training data. We can reduce this variation and improve the confidence in the calculated results by using more estimates

of the concentration profiles during the training of the neural network.

In Sections 3, 4 and 5, we have used feedforward neural networks to estimate macroscopic chemotactic sensitivity $\chi(S)$ using data calculated by relatively short simulations of Microscopic Models I, II and III, respectively. While the parabolic chemotaxis equation (2.1) provides a good description of macroscopic dynamics corresponding to Microscopic Model I on such short time scales, this is not the case when more complex Microscopic Models II and III are used. The convergence of Microscopic Model I to the parabolic chemotaxis equation (2.1) is illustrated in Figure 2(a), while we can see large quantitative differences in transient dynamics between the parabolic chemotaxis equation (2.1) and Microscopic Model II on short time scales in Figure 4(a). Consequently, the neural network estimation with loss function $\mathcal{L}(S)$ given by (2.10) works well in the case of Microscopic Model I, but we need to take into account different transient dynamics of velocity-jump processes and use loss function $\mathcal{L}_2(S)$ given by (4.16) when working with short-time simulations of Microscopic Models II and III.

Figures 4(b) and 5(b) illustrate that the loss function (4.16) works well for short simulation time windows, with the best performance obtained for $\tau = 10^{-2} = 10^2 \Delta t$. If we consider shorter time windows, then the results are influenced by noise stemming from the estimation of densities in simulations of a finite number of individuals, N , as we can see for $\tau = 10^{-3} = 10 \Delta t$ in Figures 4(b) and 5(b). On the other hand, if τ is chosen larger than optimal, then the time derivatives in the loss function (4.16) will be approximated with a bias leading to the error seen for $\tau = 10^{-1} = 10^3 \Delta t$ in Figures 4(b) and 5(b). The same conclusion can be made for Microscopic Model I in Figure 2(b) with the best performing value of the time window being $\tau = 10^{-3} = 10 \Delta t$. This is shorter than for velocity-jump processes because the transient dynamics of the parabolic chemotaxis equation (2.1) is faster than the hyperbolic chemotaxis equation (4.9) for our parameters and initial conditions. The solutions of macroscopic PDEs (2.1) and (4.9) converge as $t \rightarrow \infty$ to the same steady state solution (2.8). In particular, if a model is sufficiently simple that long-time equilibrium simulations of an individual-based model are computationally feasible, then they could also be used for estimating the chemotactic sensitivity $\chi(S)$.

Declarations.

Authorship and Contributorship: The author confirms sole responsibility for the conception, design, and execution of the study, and the preparation of the manuscript.

Conflicts of Interest: The author declares there are no conflicts of interest.

Data & Code Availability: In compliance with EPSRC's open access initiative, the data in this paper is available from <http://dx.doi.org/10.5287/ora-xqdkj4xx7>

Funding: This work was supported by the Engineering and Physical Sciences Research Council, grant number EP/V047469/1.

Ethical Approval: Not applicable.

REFERENCES

- [1] J. ADLER, *Chemotaxis in bacteria*, Science, 153 (1966), pp. 708–716.
- [2] J. AHN, *Global well-posedness and asymptotic stabilization for chemotaxis system with signal-dependent sensitivity*, Journal of Differential Equations, 266 (2019), pp. 6866–6904.
- [3] G. ARUMUGAM AND J. TYAGI, *Keller-Segel chemotaxis models: a review*, Acta Applicandae Mathematicae, 171 (2021), p. 6.
- [4] R. BAKER, C. YATES, AND R. ERBAN, *From microscopic to macroscopic descriptions of cell migration on growing domains*, Bulletin of Mathematical Biology, 72 (2010), pp. 719–762.
- [5] N. BARKAI AND S. LEIBLER, *Robustness in simple biochemical networks*, Nature, 387 (1997), pp. 913–917.
- [6] N. BELLOMO, A. BELLOUQUID, Y. TAO, AND M. WINKLER, *Toward a mathematical theory of Keller-Segel models of pattern formation in biological tissues*, Mathematical Models and Methods in Applied Sciences, 25 (2015), pp. 1663–1763.
- [7] J. DALLON AND H. OTHMER, *A discrete cell model with adaptive signalling for aggregation of Dictyostelium discoideum*, Philosophical Transactions of the Royal Society B: Biological Sciences, 352 (1997), pp. 391–417.
- [8] R. ERBAN AND S. J. CHAPMAN, *Stochastic Modelling of Reaction-Diffusion Processes*, Cambridge University Press, 2020, <https://doi.org/10.1017/9781108628389>.
- [9] R. ERBAN AND H. HWANG, *Global existence results for the complex hyperbolic models of bacterial chemotaxis*, Discrete and Continuous Dynamical Systems Series B, 6 (2006), pp. 1239–1260.
- [10] R. ERBAN, I. KEVREKIDIS, AND H. OTHMER, *An equation-free computational approach for extracting population-level behavior from individual-based models of biological dispersal*, Physica D, 215 (2006), pp. 1–24.
- [11] R. ERBAN AND H. OTHMER, *From individual to collective behaviour in bacterial chemotaxis*, SIAM Journal on Applied Mathematics, 65 (2004), pp. 361–391.
- [12] R. ERBAN AND H. OTHMER, *From signal transduction to spatial pattern formation in E. coli: A paradigm for multi-scale modeling in biology*, Multiscale Modeling and Simulation, 3 (2005), pp. 362–394.
- [13] R. ERBAN AND H. OTHMER, *Taxis equations for amoeboid cells*, Journal of Mathematical Biology, 54 (2007), pp. 847–885.
- [14] F. FILBET, P. LAURENÇOT, AND B. PERTHAME, *Derivation of hyperbolic models for chemosensitive movement*, Journal of Mathematical Biology, 50 (2005), pp. 189–207.
- [15] P. FISHER, R. MERKL, AND G. GERISCH, *Quantitative analysis of cell motility and chemotaxis in Dictyostelium discoideum by using an image processing system and a novel chemotaxis chamber providing stationary chemical gradients*, Journal of Cell Biology, 108 (1989), pp. 973–984.
- [16] B. FRANZ AND R. ERBAN, *Hybrid modelling of individual movement and collective behaviour*, in Dispersal, individual movement and spatial ecology: A mathematical perspective, M. Lewis, P. Maini, and S. Petrovskii, eds., Springer, 2012.
- [17] B. FRANZ, J. P. TAYLOR-KING, C. YATES, AND R. ERBAN, *Hard-sphere interactions in velocity-jump models*, Physical Review E, 94 (2016), p. 012129.
- [18] B. FRANZ, C. XUE, K. PAINTER, AND R. ERBAN, *Travelling waves in hybrid chemotaxis models*, Bulletin of Mathematical Biology, 76 (2014), pp. 377–400.
- [19] K. FUJIE AND T. SENBA, *Global existence and boundedness in a parabolic-elliptic Keller-Segel system with general sensitivity*, Discrete & Continuous Dynamical Systems-Series B, 21 (2016).
- [20] K. FUJIE AND T. SENBA, *Global existence and boundedness of radial solutions to a two dimensional fully parabolic chemotaxis system with general sensitivity*, Nonlinearity, 29 (2016), pp. 2417–2450.
- [21] K. FUJIE, M. WINKLER, AND T. YOKOTA, *Boundedness of solutions to parabolic-elliptic Keller-Segel systems with signal-dependent sensitivity*, Mathematical Methods in the Applied Sciences, 38 (2015), pp. 1212–1224.
- [22] R. GINIÛNAITÉ, R. BAKER, P. KULESA, AND P. MAINI, *Modelling collective cell migration: neural crest as a model paradigm*, Journal of Mathematical Biology, 80 (2020), pp. 481–504.
- [23] I. GOODFELLOW, Y. BENGIO, AND A. COURVILLE, *Deep Learning*, MIT Press, 2016.
- [24] T. HILLEN, *Transport equations and chemosensitive movement*, habilitation thesis, University of Tuebingen, Germany, 2001.
- [25] T. HILLEN, *Hyperbolic models for chemosensitive movement*, Mathematical Models and Methods in Applied Sciences, 12 (2002), pp. 1–28.
- [26] T. HILLEN AND H. OTHMER, *The diffusion limit of transport equation derived from velocity-*

- jump processes*, SIAM Journal on Applied Mathematics, 61 (2000), pp. 751–775.
- [27] T. HILLEN AND K. PAINTER, *A user's guide to PDE models for chemotaxis*, Journal of Mathematical Biology, 58 (2009), pp. 183–217.
- [28] T. HILLEN AND A. STEVENS, *Hyperbolic models for chemotaxis in 1-D*, Nonlinear Analysis: Real World Applications, 1 (2000), pp. 409–433.
- [29] R. HOCKNEY AND J. EASTWOOD, *Computer simulation using particles*, IOP Publishing Ltd, 2021.
- [30] D. HORSTMANN, *From 1970 until present: the Keller-Segel model in chemotaxis and its consequences I*, Jahresbericht der Deutschen Mathematiker-Vereinigung, 105 (2003), pp. 103–165.
- [31] D. HORSTMANN, *From 1970 until present: the Keller-Segel model in chemotaxis and its consequences. II*, Jahresbericht der Deutschen Mathematiker-Vereinigung, 106 (2004), pp. 51–69.
- [32] E. KELLER AND L. SEGEL, *Traveling bands of chemotactic bacteria: A theoretical analysis*, Journal of Theoretical Biology, 30 (1971), pp. 235–248.
- [33] A. KRAUSE, E. GAFFNEY, AND B. WALKER, *Concentration-dependent domain evolution in reaction–diffusion systems*, Bulletin of Mathematical Biology, 85 (2023), p. 14.
- [34] S. LEE, Y. PSARELLIS, C. SIETOS, AND I. KEVREKIDIS, *Learning black-and gray-box chemotactic PDEs/closures from agent based Monte Carlo simulation data*, Journal of Mathematical Biology, 87 (2023), p. 15.
- [35] T. LORENZI, K. PAINTER, AND C. VILLA, *Phenotype structuring in collective cell migration: a tutorial of mathematical models and methods*, Journal of Mathematical Biology, 90 (2025), p. 61.
- [36] G. MACDONALD, J. MACKENZIE, M. NOLAN, AND R. INSALL, *A computational method for the coupled solution of reaction–diffusion equations on evolving domains and manifolds: Application to a model of cell migration and chemotaxis*, Journal of Computational Physics, 309 (2016), pp. 207–226.
- [37] M. MAYO AND R. SOTO, *Bacterial chemotaxis considering memory effects: Derivation of the reaction-diffusion equations*, Physical Review E, 111 (2025), p. 054409.
- [38] A. MUINONEN-MARTIN, D. VELTMAN, G. KALNA, AND R. INSALL, *An improved chamber for direct visualisation of chemotaxis*, PLoS One, 5 (2010), p. e15309.
- [39] H. OTHMER, S. DUNBAR, AND W. ALT, *Models of dispersal in biological systems*, Journal of Mathematical Biology, 26 (1988), pp. 263–298.
- [40] H. OTHMER AND P. SCHAAP, *Oscillatory cAMP signaling in the development of Dictyostelium discoideum*, Comments on Theoretical Biology, 5 (1998), pp. 175–282.
- [41] H. OTHMER AND A. STEVENS, *Aggregation, blow up and collapse: The ABC's of taxis in reinforced random walks*, SIAM Journal on Applied Mathematics, 57 (1997), pp. 1044–1081.
- [42] C. PATLAK, *Random walk with persistence and external bias*, Bulletin of Mathematical Biophysics, 15 (1953), pp. 311–338.
- [43] Y. PSARELLIS, S. LEE, T. BHATTACHARJEE, S. DATTA, J. BELLO-RIVAS, AND I. KEVREKIDIS, *Data-driven discovery of chemotactic migration of bacteria via coordinate-invariant machine learning*, BMC Bioinformatics, 25 (2024), p. 337.
- [44] P. SPIRO, J. PARKINSON, AND H. OTHMER, *A model of excitation and adaptation in bacterial chemotaxis*, Proceedings of the National Academy of Sciences USA, 94 (1997), pp. 7263–7268.
- [45] J. TAYLOR-KING, B. FRANZ, C. YATES, AND R. ERBAN, *Mathematical modelling of turning delays in swarm robotics*, IMA Journal of Applied Mathematics, 80 (2015), pp. 1454–1474.
- [46] A. WAITE, N. FRANKEL, AND T. EMONET, *Behavioral variability and phenotypic diversity in bacterial chemotaxis*, Annual Review of Biophysics, 47 (2018), pp. 595–616.
- [47] M. WINKLER, *Finite-time blow-up in the higher-dimensional parabolic-parabolic Keller-Segel system*, Journal de Mathématiques Pures et Appliquées, 100 (2013), pp. 748–767.
- [48] M. WINKLER, *Unlimited growth in logarithmic Keller-Segel systems*, Journal of Differential Equations, 309 (2022), pp. 74–97.
- [49] C. XUE, *Macroscopic equations for bacterial chemotaxis: integration of detailed biochemistry of cell signaling*, Journal of Mathematical Biology, 70 (2015), pp. 1–44.
- [50] S. YASUDA, *Effects of internal dynamics on chemotactic aggregation of bacteria*, Physical Biology, 18 (2021), p. 066001.
- [51] J. ZHANG AND R. ERBAN, *From Brownian dynamics to Poisson-Nernst-Planck equations: multi-resolution simulations of ions*, arXiv preprint arXiv:2506.19738, (2025).
- [52] W. ZHAO, H. ZHAO, M. LI, AND C. HUANG, *Microfluidic devices for neutrophil chemotaxis studies*, Journal of Translational Medicine, 18 (2020), p. 168.