

Bayesian inference on the Keller–Segel model

Thomas Goodwin¹ Christian Evenhuis²
Stephen Woodcock³ Matias Quiroz⁴

(Received 28 February 2020; revised 26 June 2020)

Abstract

The Keller–Segel (KS) model is a system of partial differential equations that describe chemotaxis—how cells move in response to chemical stimulus. Simulated data in the form of cell counts are used to carry out Bayesian inference on the KS model. A Bayesian analysis on the KS model is performed on three sets of initial conditions. First, the KS model is solved numerically using a finite difference method and Bayesian inference is performed on parameters of the model such as the *cell diffusion* and chemical sensitivity. We investigate the predictive posterior distribution of future data and the convergence of the 95% credible interval of cell diffusion at different grid sizes using the three different initial conditions.

DOI:10.21914/anziamj.v61i0.15185, © Austral. Mathematical Soc. 2020. Published 2020-08-10, as part of the Proceedings of the 14th Biennial Engineering Mathematics and Applications Conference. ISSN 1445-8810. (Print two pages per sheet of paper.) Copies of this article must not be made otherwise available on the internet; instead link directly to the DOI for this article.

Contents

1	Introduction	C182
1.1	Data	C184
2	Numerical scheme for the Keller–Segel model	C184
2.1	Boundary conditions	C186
2.2	Initial conditions	C187
3	Bayesian inference on the Keller–Segel model	C188
3.1	In context of PDEs	C189
3.2	Ensemble sampler	C189
3.3	Likelihood	C189
3.4	Simulation setting and prior selection	C190
4	Results	C191
5	Discussion and conclusion	C192

1 Introduction

Chemotaxis is the movement of an organism in response to a chemical stimulus [3]. The most well-known and heavily studied mathematical model of chemotaxis is the *Keller–Segel Model* (KS) [8] which was first introduced in 1970. In the KS model the cells undergo a random walk which, at a population level, corresponds to a diffusion process. In the model, the response of the cells to chemical stimulus is modelled by biasing the random walk in the direction towards or away from the *gradient* of the chemical concentration. This gives a particular reaction-diffusion system of partial differential equations. Modeling chemotaxis from a differential equation point of view provides a population-based model of cells which does not require any cells to be distinguished, but rather assumes a homogeneous population density of cells.

The generalized Keller–Segel Model for chemotaxis is

$$\begin{aligned} u_t &= \nabla[k_1(u, v)\nabla u + k_2(u, v)u\nabla v] + k_3(u, v), \\ v_t &= D_v\Delta v + k_4(u, v) - k_5(u, v)v, \end{aligned} \quad (1)$$

where $u(x, t)$ denotes the cell density at position x at time t , $v(x, t)$ denotes the chemical signal concentration, k_1 is the diffusivity of the cells, k_2 is the chemo-tactic sensitivity, k_3 describes cell growth and death, k_4 represents the chemo-repellent production and k_5 the chemo-repellent degradation. As shown by equation (1), the cell migration is dependent on the chemical gradient. Here, the chemicals' primary mechanism of movement follows Fick's second law; that is, the diffusion equation which describes how diffusion changes a chemicals concentration over time. The cells' movement is described with a diffusion term plus the attraction to or repulsion from the chemical. We set k_3 equal to zero as we assume there is no change in the number of cells over time. Furthermore, k_4 and k_5 are also set equal to zero, which corresponds to no production or degradation of the chemo-repellent.

The main focus of this article is performing Bayesian inference on the KS model. The main goal of Bayesian inference on the KS model is to derive the posterior distribution over the parameter space Θ given the data. Sherratt et al. [10] fitted the KS model to experimental data of chemically controlled cell movement. One drawback of the KS model discussed by Sherratt et al. is the specification of the coefficients of the model with the chemical concentration. Performing Bayesian inference eliminates the need to specify specific values of the parameter vector $\theta \in \Theta$ when modelling data, but instead estimates them via the probability distribution of θ given the data. This so-called posterior distribution is obtained by specifying a prior distribution and a likelihood function for the observed data. Furthermore, as a prior distribution, we incorporate work done by Harvath and Aksamit [7] on the chemo-attractant activity and migration of neutrophil towards certain peptides.

1.1 Data

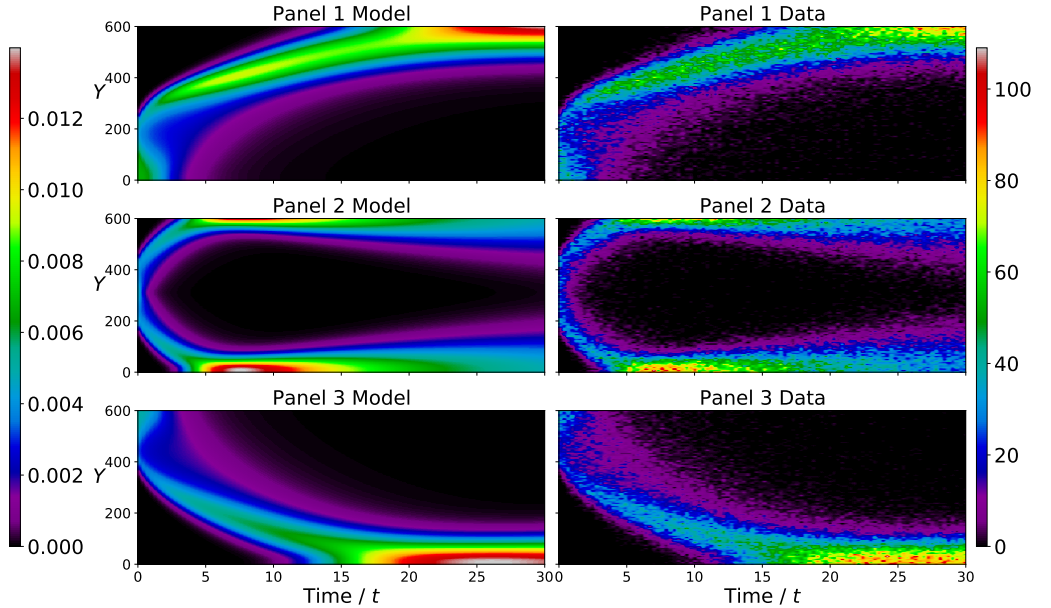
We demonstrate our approach using simulated data sets. This has the advantage of better assessing the method since we can empirically study its convergence properties; for example, that we recover the true parameters as we increase the number of data points, and that the predictive inference is able to correctly forecast the out-of-sample data.

The data is in the form of $M \times N$ arrays which are cell counts at each of the M discretized spatial steps conditional on each of the N time points. Intuitively, these arrays are the distribution of cells throughout the spatial domain y conditional on time t . A time grid array (t -grid) of $N + 1$ points represents the small time steps at which the data was recorded, from the starting time t_0 to the final time t_f . The spatial grid array (y -grid) represents the domain and consists of M spatial steps over the interval $[0, 600]$. Three trials, referred to here as Panels, corresponding to different initial conditions of the cells at $t = 0$. The simulated data was generated by first solving the KS model numerically and then, after multiplying by the number of cells to get the data in terms of cell counts, simulating Poisson random variables with expected value equal to the solution of the KS model at each spatial grid point. Applying a Poisson random variable gives the data in terms of the number of cells for each spatial grid point in the domain over time t . The initial conditions of the Panels are discussed in Section 2. The three different Panels are pictured in Figure 1.

2 Numerical scheme for the Keller–Segel model

The Backward Time Central Space (BTCS) implicit method is used to solve the KS model numerically. The backward method was used due to its robustness at different step sizes; however, one disadvantage is that it is computationally expensive due to the inverse of the tri-diagonal matrix required when solving a discretized version of the KS model. In one dimension, the KS *minimal*

Figure 1: Model (left) and simulated data (right). The model data shows areas of higher probability over spatial domain y and time t , coloured according to the left colour-map. The simulation data shows the distribution of cells across y and t , coloured according to the right colour-map.



model is

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial}{\partial x} \left(\alpha \frac{\partial u}{\partial x} - \chi u \frac{\partial v}{\partial x} \right), \quad t \in [0, T], \quad x \in [0, +L], \\ \frac{\partial v}{\partial t} &= c \frac{\partial^2 v}{\partial x^2}, \end{aligned} \tag{2}$$

where α and c are the diffusion coefficients of the cells and chemical, respectively, χ is the chemo-sensitivity term, and L is the length of the system. Different forms of the sensitivity parameter χ exist in the literature. Two forms suggested are a positive constant χ_1 [1] and a logarithmic law, $\chi \propto 1/c$ [2, 4]. We define the chemotactic sensitivity function χ , a variant of the receptor ki-

netic law function proposed by Mesibov et al. [9]: $\chi \frac{\partial v}{\partial x} = \chi_0 V_{\max} / (V_{\max} + |\frac{\partial v}{\partial x}|)$, where χ_0 is the chemo-sensitivity coefficient and V_{\max} is the maximum velocity of the cells. This velocity restricts the speed of the cells as they become saturated with the chemical repellent.

Let U_i^n denote the discretized value of $u(x_i, t_n)$ at points i, n . The discretization of the first line of equation (2) is

$$\frac{U_i^n - U_i^{n-1}}{\Delta t} = \alpha \frac{U_{i+1}^n - 2U_i^n + U_{i-1}^n}{(\Delta x)^2} - \frac{\phi_{i+1}^n U_{i+1}^n - \phi_{i-1}^n U_{i-1}^n}{2\Delta x}, \quad (3)$$

where $\phi = \chi \frac{\partial v}{\partial x}$, Δt is the time step, and Δx is the spatial step. In the above equation, the two functions in the term $\chi \frac{\partial v}{\partial x}$ are coupled, and so the first order central difference is approximated as one function instead of using the product rule. The substitution of ϕ makes equation (3) much more general and it can serve as a numerical solver for a more general class of advection-diffusion equations. Performing the coupled discretization also makes the derivation more elegant. Solving for U_i^{n-1} creates the tri-diagonal matrix with each term on the diagonals

$$U_i^{n-1} = -(A + \phi_{i-1}^n)U_{i-1}^n + (2A + 1)U_i^n - (A - \phi_{i+1}^n)U_{i+1}^n,$$

with constant $A = \frac{\alpha \Delta t}{(\Delta x)^2}$.

The well-known discretization of the second line of equation (2) (also known as the heat equation) for the chemical $v(x, t)$, using the BTCS implicit method can be shown to construct the tri-diagonal matrix

$$V_i^{n-1} = -rV_{i-1}^n + (1 + 2r)V_i^n - rV_{i+1}^n,$$

where $r = \frac{\alpha \Delta t}{(\Delta x)^2}$.

2.1 Boundary conditions

The cells or chemical cannot move through the boundaries, hence, no-flux boundary conditions are necessary to capture this effect. The analytical flux

of the cells is defined by

$$J_u = \alpha \frac{\partial u}{\partial x} - \chi \frac{\partial v}{\partial x} u.$$

At the boundaries $[0, L]$, with the flux equal to zero, we obtain a *Robin* type boundary condition, which is a weighted linear combination of Dirichlet and Neumann boundary conditions [6]. Practically, on a discretized domain, flux is analysed at the boundaries with the addition of a spatially changing advection term.

The left and right boundary conditions are obtained from (3) for $i = 0$ and $i = 50$, respectively. But these boundary conditions create so-called ghost nodes at points $i = -1$ and $i = 51$ which are resolved by setting $U_{-1}^n = U_0^n$, $\phi_{-1}^n = \phi_0^n$, $U_{51}^n = U_{50}^n$ and $\phi_{51}^n = \phi_{50}^n$. The resulting boundary conditions are

$$\begin{aligned} \text{Left Boundary: } U_0^{n-1} &= (A - \phi_0^n + 1)U_0^n - (A - \phi_1^n)U_1^n, \\ \text{Right Boundary: } U_{50}^{n-1} &= -(A + \phi_{49}^n)U_{49}^n + (A + \phi_{50}^n + 1)U_{50}^n. \end{aligned}$$

To satisfy the conservation of mass property of the chemical, the boundary conditions of the chemical are

$$\begin{aligned} \text{Left Boundary: } V_0^{n-1} &= (1 + r)V_0^n + (-r)V_1^n, \\ \text{Right Boundary: } V_{50}^{n-1} &= (-r)V_{49}^n + (1 + r)V_{50}^n, \end{aligned}$$

because the homogeneous Neumann boundary conditions given by spatial derivatives $v_x(L, t) = 0$ and $v_x(0, t) = 0$ give discrete values $V_{51}^n = V_{50}^n$ and $V_{-1}^n = V_0^n$.

2.2 Initial conditions

The initial conditions are constructed by dividing the spatial dimension into three equal regions and setting the concentration of cells and chemical in each to be constant. The initial conditions are similar to those described by

Table 1: Initial conditions of KS model.

	Panel 1	Panel 2	Panel 3
Cells	[1, 0, 0]	[0, 1, 0]	[0, 0, 1]
Chemical	[1, 0, 0.25]	[0, 1, 0]	[0, 0, 1]

Sherratt et al. [10] for the Boyden chamber assay except we model a third partition, and thus have a lower, upper and middle well, into which we are free to inject cells and the chemo-repellant at the initial time. The three divisions of the spatial domain $[0, 600]$, referred to as lanes, are

Lane 1 : $0 \leq x < 200$, **Lane 2** : $200 \leq x < 400$, **Lane 3** : $400 \leq x < 600$.

By assigning a concentration value to each of the lanes, we define uniform distributions of the concentration of chemical $v(x, t)$ or cells $u(x, t)$ injected at the initial time. For convenience the initial conditions are represented as lists of the chemical and cell concentrations in each lane. Panel 1 has $[1, 0, 0]$ and $[1, 0, 0.25]$ which represents 100% of the cells and 100% of the chemical injected into Lane 1 and 25% of the chemical injected into Lane 3. The edges of the initial conditions were also slightly smoothed to reduce error in the numerical solver.

3 Bayesian inference on the Keller–Segel model

We use Bayesian inference to find the posterior probability densities of the parameters in the model. This work is an ideal application of Bayesian inference: there is a generative model parameterised by variables $\theta \in \Theta$, we establish priors to limit the range of Θ , and seek a probability distribution for θ which indicates the subset of Θ compatible with the observed data.

3.1 In context of PDEs

Computing expectations with respect to the posterior distribution in Bayesian analysis requires computing complicated integrals. The integrals can be estimated by simulation if we are able to sample from the posterior distribution of the parameters. Often, thousands of samples are taken to obtain an accurate representation. Evaluating the likelihood in our context involves numerically solving the system of partial differential equations for each required posterior sample.

3.2 Ensemble sampler

Goodman and Weare in 2010 [5] proposed the ensemble sampler, a Monte Carlo Markov Chain (MCMC) algorithm, to sample from distributions. The MCMC is used to sample probability distribution, particularly when direct sampling or other simpler sampling schemes are not possible. One of the benefits of the ensemble sampler is that it is *affine invariant*; that is, unchanged by linear transformations and translations of the parameters. This allows sampling from skewed distributions to be significantly faster than standard MCMC methods which may require coordinate re-scaling or hard to compute determinants for the transformations.

3.3 Likelihood

Assuming conditional independence of the data given the parameters, the log-likelihood of observed outcome $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$ is

$$\log \mathcal{L}(\Theta \mid \mathbf{X}) = \sum_{i=1}^N \log p(\mathbf{x}_i \mid \Theta),$$

for probability density $p(\cdot \mid \cdot)$.

Our data are cell counts per spatial interval, which we model as Poisson data

with log-likelihood

$$\log \mathcal{L}(\theta \mid \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = -n\theta - \sum_{i=1}^n \log(x_i!) + \log(\theta) \sum_{i=1}^n x_i.$$

3.4 Simulation setting and prior selection

The parameters we subject to inferencing are the three coefficients in (2) and the number of cells in the domain $[\alpha, \chi_0, c, \text{ncells}]$. We introduce a small time offset parameter $\hat{t}_0 \in (0, 0.5]$ to the model which delays the start time by t_0 . Also, a spatial offset parameter $\hat{y}_0 \in [-15, 15]$ shifts the boundaries by that value. The addition of these parameters take into account experimental errors that arise in practical situations. To implement these constraints, the model is simulated on the domain $t \in [2, 30]$, $y = [30, 570]$.

We set the true parameters to be in the ranges of estimates determined by Tranquillo et al. [11]. Defining the spatial grid in micrometers (μm), this results in $\theta_{\text{true}} = [\alpha, c, \chi_0, \text{ncells}, \hat{t}_0, \hat{y}_0] = [300, 2000, 6 \times 10^6, 1000, 0.42, 10]$ which are the underlying parameters of the data simulated from the model.

For each of the three Panels, a Bayesian model was constructed with six parameters and their priors. We proposed, $\theta_{\text{prop}} = [800, 1900, 3 \times 10^6, 900, 0.3, 5]$ which serves as the initial value of the Nealder–Mead algorithm to find the maximum *a posteriori* (MAP) estimation. In this case, uninformative, or flat, priors were used on the first four parameters to get an idea of the posterior densities as more evidence becomes available. Priors are also enforced with a positivity constraint on parameters such as the diffusion coefficients and time offset. We sample 14000 draws from the posterior and use the first 4000 samples as burn-in.

Table 2: Prior distribution of parameters.

Parameter	Prior Distribution
Cell diffusion coefficient α	$\text{lognorm}(1648.82, 2161.2^2)$
Chemo-repellent diffusion coefficient c	$\text{lognorm}(3662.6, 5618.67^2)$
Chemo-tactic sensitivity χ_0	$\text{lognorm}(8243\,706, 10805987^2)$
Number of cells	$\text{lognorm}(1877.13, 1303.9^2)$
Time offset \hat{t}_0	$\text{beta}(1.0, 1.0)$
Spatial \hat{y}_0	$N(0, 10^2)$

4 Results

To predict a future observation \tilde{y} at, say, time step $t + 1$, the predictive posterior distribution is computed from

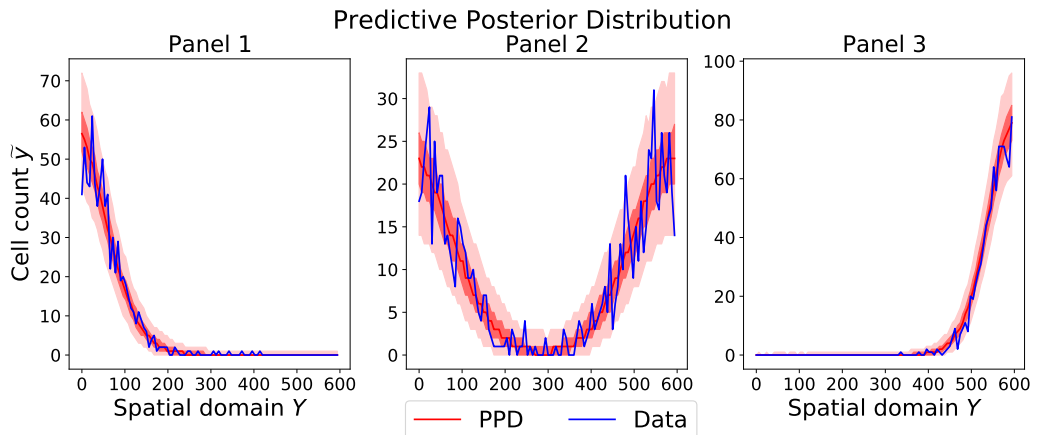
$$p(\tilde{y} \mid y) = \int_{\Theta} p(\tilde{y}, \theta \mid y) \, d\theta = \int_{\Theta} p(\tilde{y} \mid \theta) p(\theta \mid y) \, d\theta. \tag{4}$$

Simulated data is generated on a $(M, N + 1)$ grid and the posterior model $p(\theta \mid y)$ is obtained withholding the last observation. Based on this posterior, we easily simulate $p(\tilde{y} \mid y)$ from (4) and compare the predictive distribution to the out-of-sample true cell counts at the last time-point.

Figure 2 shows the distribution of future observations of the data in Panels 1, 2 and 3 compared to the simulated data. Using three different initial conditions, this figure shows that our posterior distribution gave a good prediction at future data points.

To test the convergence of the Bayesian posterior to the true parameters θ_{true} as more data becomes available, the posterior was sampled for each Panel at different grid sizes. The model was tested on $M \times N$ square grids ($M = N$) for $M = 50, 75, 100, 200, 500$. The 95% credible intervals are plotted for cell diffusion α for each Panel at each grid size. Figure 3 shows the convergence of the Bayesian posterior to the true parameter α at varying grid sizes. Theoretically, by the Bernstein–von Mises theorem [12, Chapter 10.2], we

Figure 2: Predictive posterior distribution (PPD) versus simulated data for Panels 1, 2, and 3.



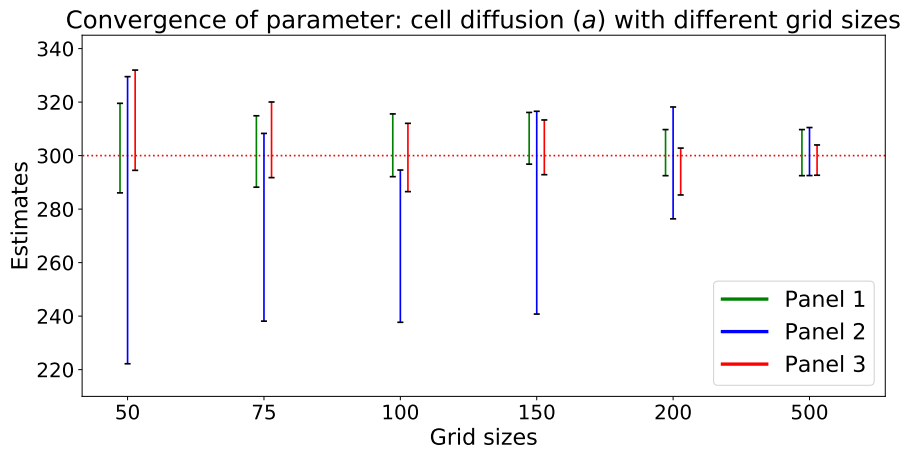
expect the scale of the posterior to asymptotically be $\propto 1/\sqrt{n}$, where n is the number of data points.

The convergence of cell diffusion α in each Panel, shown in Figure 3, gives insight into experimental conditions because different initial conditions correspond to different experimental conditions. We see in Figure 3 that at smaller grid sizes, Panel 2 poorly estimates parameter α , while Panels 1 and 3 give more accurate estimates of the parameter on all grid sizes. Convergence analysis was done for all parameters (not shown) which yield similar results as cell diffusion α .

5 Discussion and conclusion

We demonstrated how to carry out Bayesian inference in the Keller–Segel model. We computed the posterior predictive distribution and, moreover, demonstrated empirical convergence of the model parameters as the grid size increased in a simulated data setting. The simulated data was obtained using established parameter values in the literature. By analyzing the convergence

Figure 3: Convergence of α to \mathbf{a} for varying grid sizes.



of these parameters to their true values, we find that at moderate square grid sizes (50, 100, 150), certain initial conditions give better estimates of certain parameters. For example, as seen by Figure 3, Panel 3 gives good estimates of the cell diffusion α at all grid-sizes, in contrast to Panel 2. By determining which Panel corresponds to more accurate estimates of each parameter, we can inform the experimenters on what experimental condition to use with what parameter they seek to determine. As a way to obtain more accurate credible intervals for the parameters, future work will include *hierarchical models* which give a way of using the posterior distribution as a prior distribution in the next iteration of sampling from the posterior. Future work will also include reducing the number of time grid points to determine the trade-off between accurate estimates and computation time, as well as the addition of a background noise parameter which accounts for the experimental error of false detection of cells.

References

- [1] D. Balding and D. L. S. McElwain. “A mathematical model of tumour-induced capillary growth”. In: *J. Theor. Biol.* 114.1 (1985), pp. 53–73. DOI: [10.1016/S0022-5193\(85\)80255-1](https://doi.org/10.1016/S0022-5193(85)80255-1) (cit. on p. [C185](#)).
- [2] D. A. Brown and H. C. Berg. “Temporal stimulation of chemotaxis in *Escherichia coli*”. In: *Proc. Nat. Acad. Sci.* 71.4 (1974), pp. 1388–1392. DOI: [10.1073/pnas.71.4.1388](https://doi.org/10.1073/pnas.71.4.1388) (cit. on p. [C185](#)).
- [3] H. Chisholm. *The Encyclopædia britannica: a dictionary of arts, sciences, literature and general information*. Vol. 6. Encyclopaedia Britannica Co., 1910 (cit. on p. [C182](#)).
- [4] F. W. Dahlquist, P. Lovely, and D. E. Koshland. “Quantitative analysis of bacterial migration in chemotaxis”. In: *Nature New Biol.* 236.65 (1972), pp. 120–123. DOI: [10.1038/newbio236120a0](https://doi.org/10.1038/newbio236120a0) (cit. on p. [C185](#)).
- [5] J. Goodman and J. Weare. “Ensemble samplers with affine invariance”. In: *Commun. Appl. Math. Comput. Sci.* 5.1 (2010), pp. 65–80. URL: <https://projecteuclid.org/euclid.camcos/1513731992> (cit. on p. [C189](#)).
- [6] K. Gustafson and T. Abe. “The third boundary condition—was it Robin’s?” In: *Math. Intell.* 20.1 (1998), pp. 63–71. DOI: [10.1007/BF03024402](https://doi.org/10.1007/BF03024402) (cit. on p. [C187](#)).
- [7] L. Harvath and R. R. Aksamit. “Oxidized N-formylmethionyl-leucyl-phenylalanine: Effect on the activation of human monocyte and neutrophil chemotaxis and superoxide production”. In: *J. Immun.* 133.3 (1984), pp. 1471–1476. URL: <https://www.jimmunol.org/content/133/3/1471> (cit. on p. [C183](#)).
- [8] E. F. Keller and L. A. Segel. “Initiation of slime mold aggregation viewed as an instability”. In: *J. Theor. Bio.* 26.3 (1970), pp. 399–415. DOI: [10.1016/0022-5193\(70\)90092-5](https://doi.org/10.1016/0022-5193(70)90092-5) (cit. on p. [C182](#)).

- [9] R. Mesibov, G. W. Ordal, and J. Adler. “The range of attractant concentrations for bacterial chemotaxis and the threshold and size of response over this range: Weber law and related phenomena”. In: *J. Gen. Physiol.* 62.2 (1973), pp. 203–223. DOI: [10.1085/jgp.62.2.203](https://doi.org/10.1085/jgp.62.2.203) (cit. on p. [C186](#)).
- [10] J. A. Sherratt, E. H. Sage, and J. D. Murray. “Chemical control of eukaryotic cell movement: A new model”. In: *J. Theor. Biol.* 162.1 (1993), pp. 23–40. DOI: [10.1006/jtbi.1993.1074](https://doi.org/10.1006/jtbi.1993.1074) (cit. on pp. [C183](#), [C188](#)).
- [11] R. T. Tranquillo, S. H. Zigmond, and D. A. Lauffenburger. “Measurement of the chemotaxis coefficient for human neutrophils in the under-agarose migration assay”. In: *Cell Motil. Cytoskel.* 11.1 (1988), pp. 1–15. DOI: [10.1002/cm.970110102](https://doi.org/10.1002/cm.970110102) (cit. on p. [C190](#)).
- [12] A. W. van der Vaart. *Asymptotic Statistics*. Vol. 3. Cambridge University Press, 2000. DOI: [10.1017/CB09780511802256](https://doi.org/10.1017/CB09780511802256) (cit. on p. [C191](#)).

Author addresses

1. **Thomas Goodwin**, School of Mathematical and Physical Sciences, University of Technology, Sydney 2007, AUSTRALIA.
<mailto:thomas.goodwin@uts.edu.au>
orcid:[0000-0003-1394](https://orcid.org/0000-0003-1394)
2. **Christian Evenhuis**, School of Mathematical and Physical Sciences, University of Technology, Sydney 2007, AUSTRALIA.
<mailto:christian.evenhuis@uts.edu.au>
3. **Stephen Woodcock**, School of Mathematical and Physical Sciences, University of Technology, Sydney 2007, AUSTRALIA.
<mailto:stephen.woodcock@uts.edu.au>
4. **Matias Quiroz**, School of Mathematical and Physical Sciences, University of Technology, Sydney 2007, AUSTRALIA.

<mailto:matias.quiroz@uts.edu.au>