Accepted Manuscript

The invasion speed of cell migration models with realistic cell cycle time distributions

Enrico Gavagnin, Matthew J. Ford, Richard L. Mort, Tim Rogers, Christian A. Yates

1 Highlights

- 2 • Given the average cell-cycle time, the minimum invasion speed is obtained when the cell-
- 3 cycle time is Dirac-delta distributed;

- 4 • We find the range of variability for the speed of mathematical models of cell migration
- 5 which adopt realistic hypoexponential cell-cycle time distributions;

- 6 • The maximum speed adopting hypoexponential distributions is obtained by using an ex-
- 7 ponentially distributed cell-cycle time;

- 8 • We find an analytical expression for the invasion speed of general Eralng cell-cycle time
- 9 distribu- tions.

1

The invasion speed of cell migration models with realistic cell cycle time distributions

Enrico Gavagnin[*1], Matthew J. Ford[2], Richard L. Mort[3], Tim Rogers[1], and Christian A. Yates[1]

[1] *Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, UK*

[2] *Centre for Research in Reproduction and Development, McGill University, Montréal, H3G 1Y6, Québec*

[3] *Division of Biomedical and Life Sciences, Faculty of Health and Medicine, Lancaster University, Bailrigg, Lancaster LA1 4YG, UK*

Abstract

Cell proliferation is typically incorporated into stochastic mathematical models of cell migration by assuming that cell divisions occur after an exponentially distributed waiting time. Experimental observations, however, show that this assumption is often far from the real cell cycle time distribution (CCTD). Recent studies have suggested an alternative approach to modelling cell proliferation based on a multi-stage representation of the CCTD.

In this paper we investigate the connection between the CCTD and the speed of the collective invasion. We first state a result for a general CCTD, which allows the computation of the invasion speed using the Laplace transform of the CCTD. We use this to deduce the range of speeds for the general case. We then focus on the more realistic case of multi-stage models, using both a stochastic agent-based model and a set of reaction-diffusion equations for

*Corresponding author: e.gavagnin@bath.ac.uk

2

the cells' average density. By studying the corresponding travelling wave solutions, we obtain

an analytical expression for the speed of invasion for a general N-stage model with identical

transition rates, in which case the resulting cell cycle times are Erlang distributed. We show

that, for a general N-stage model, the Erlang distribution and the exponential distribution

lead to the minimum and maximum invasion speed, respectively. This result allows us to

determine the range of possible invasion speeds in terms of the average proliferation time for

any multi-stage model.

Keywords: Cell migration, multi-stage model, cell cycle time distribution, invasion speed,

agent-based model, travelling wave.

1 Introduction

Cellular invasion is a process of fundamental importance in numerous morphogenetic and patho-

logical mechanisms. Important examples of processes in which cell migration plays a crucial role

include embryonic development [Gilbert, 2003, Keller, 2005], wound healing [Maini et al., 2004,

Deng et al., 2006] and tumour invasion [Hanahan and Weinberg, 2000].

Understanding how the properties of the individual cells contribute to the formation and

the propagation of the invasion wave is of fundamental importance. In fact, this can reveal the

micro-scale mechanisms that are responsible for a given phenomenological aspect, and hence

suggest effective therapeutic approaches to inhibit, or enhance, cell migration by interrupting

the cell cycle [Sadeghi et al., 1998, Gray-Schopfer et al., 2007, Haass and Gabrielli, 2017].

Despite the large variety of actions and interactions which cells can undergo, there are at

least two aspects of cells' behaviour which are essential in order for the invasion to take place.

These are cell motility and cell proliferation [Simpson et al., 2007, Mort et al., 2016]. If one

of these two aspects does not occur properly, the impact on the collective invasion is typically

evident and it can affect the success of the colonisation. For example, Mort et al. [2016] show

3

53 using joint experimental and a modelling approach that the failure of colonisation of the mouse

54 embryo by melanoblasts in Kit mutants is probably driven by reduced proliferation.

55 Extensive research has focused on the effect that cell motility and proliferation behaviours

56 have on the speed of the invasion, c. The common approach makes use of simple mathematical

57 models which typically take the form of a stochastic agent-based model (ABM) [Anderson and

58 Chaplain, 1998, Deutsch and Dormann, 2007] or a deterministic partial differential equation

59 (PDE) [Murray, 2007, Wise et al., 2008]. By computing the invasion speed of the model, either

60 analytically or numerically, it is possible to link the parameters which modulate the movement

61 and proliferation with the speed of invasion.

62 Many studies have investigated this link in more general contexts, beginning with the seminal

63 work of Fisher [1937] and Kolmogorov A. [1991], and more recently more complex models of

64 populations with multiple types or stages [Elliott and Cornell, 2012, Neubert and Caswell, 2000].

65 From these studies, it is well known that, when agents' motility is modelled as diffusion with

66 diffusion coefficient D and proliferation occurs at rates λ, the invasion speed is proportional to

67 the square root of the product of the rates, i.e. $c \propto \sqrt{D\lambda}$ [Fisher, 1937].

68 It is important to notice the that majority of the literature on the speed of invasion of trav-

69 elling waves is based on the assumption that proliferation events occur as independent Poisson

70 processes [Simpson et al., 2007, Mort et al., 2016]. In the context of cell migration, this is

71 equivalent to assuming that cells proliferate after an exponentially distributed random time.

72 However, experimental observations show that the cell cycle time distribution (CCTD) is typi-

73 cally non-monotonic and differs substantially from an exponential distribution (see Figure 1 (f)

74 for an example) [Golubev, 2016, Yates et al., 2017, Chao et al., 2018].

75 There is a vast literature regarding the appropriate representation of the CCTD [Csikász-

76 Nagy et al., 2006, Gérard and Goldbeter, 2009, Powathil et al., 2012]. One class of representa-

4

tions, known as multi-stage models (MSMs), have gained particular attention in several recent studies [Yates et al., 2017, Vittadello et al., 2018, Chao et al., 2018]. The main idea of MSMs is to partition the cell cycle into N sequential stages. As time evolves, each cell can transit from one stage, i, to the next one, $i+1$, after an exponentially distributed waiting time with parameter λ_i. When a cell is found at the last stage, N, it can proliferate with rate λ_N, which leads the cell to split into two daughter cells, both initialised at the first stage. The main motivation that makes MSMs mathematically appealing is the Markov property of the exponentials which simplifies both the analytical investigation of the model and its computational implementation. Moreover, MSMs lead to CCTDs that are hypoexponential and hypoexponential distributions have been shown to provide an excellent agreement with experimental data [Golubev, 2016, Yates et al., 2017, Chao et al., 2018].

Despite the fact that there is some evidence to suggest that the cell cycle is a series of uncoupled exponentially distributed phases [Chao et al., 2018], Yates et al. [2017] were at pains to point out that the stages in their MSM do not correspond to the phases in the cell cycle, but are tools which allow to fit the correct cell cycle distribution. Similarly, here, we are reticent to link the N stages of our model to N realistic steps in biological cell cycles. Especially since, when fitting to experimental data, different choices of N can give almost equally good agreement to cell cycle distributions data.' In particular, the *stages* of the MSMs should not be confused with the biological *phases* of the cell cycle which, in general, are not exponentially distributed (see Figure 1) [Chao et al., 2018].

Whilst previous studies have investigated MSMs extensively in the case of spatially uniform scenarios [Yates et al., 2017], there is still little understanding about the effect which MSMs have on invading waves of cells. In particular, it is not clear how, and to what extent, a multi-stage representation of the CCTD can impact on the speed of invasion.
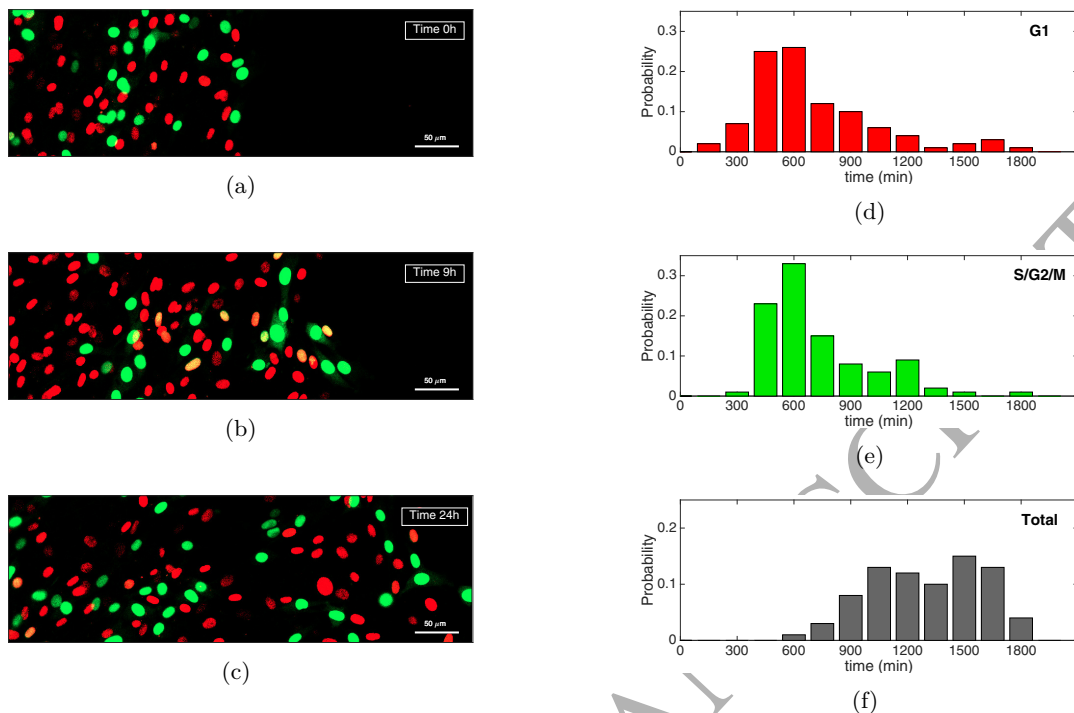
Figure 1: Panels (a-c): Mouse NIH-3T3 fibroblasts with Fucci2a status migrating into open space [Mort et al., 2014]. The Fucci2a system incorporates genetically encoded probes that highlight in red the nuclei of cells in the G1 phase and in green those of cells in one of the other phases, S/G2/M. Panels (d-f): experimental distributions of the time length of the G1 phase (panel (d)), S/G2/M phases (panel (e)) and total CCTD (panel (f)). Both the G1 and S/G2/M distributions show a clear non-monotonic trend, which indicates that are not exponentially distributed. To capture both these non-monotonicities using a MSM for the CCTD, a minimum of four stages is required, two for each of the two phases.

The most recent progress on this was made by Vittadello et al. [2018]. In their work, the authors derive an analytical expression for the invasion speed of a 2-stage MSM in terms of the two rates of stage transition, λ_1, λ_2, and the diffusion coefficient of cells, D:

$$c = \sqrt{2D \left[-\lambda_1 - \lambda_2 + \sqrt{\lambda_1^2 + 6\lambda_1\lambda_2 + \lambda_2^2} \right]}. \tag{1}$$

101 The findings of Vittadello et al. [2018] provide useful insights in the qualitative effect of the MSMs.

102 However a general expression for the invasion speed, as in equation (1), but for biologically

6

realistic MSMs, which typically have ten or more stages [Yates et al., 2017, Chao et al., 2018], is not feasible analytically. Hence, there are important questions about the quantitative effect of MSMs on the invasion speed which remain unanswered. In particular, the range of variability in speed for a general N-stage MSM has yet to be studied.

To investigate the effect of incorporating a general CCTD into the invasion models, we follow two distinct approaches. In the first part of the paper, we formulate a generalisation of the Fisher-KPP equation in which the cell population is structured by age. By studying the traveling wave solutions of the model, we derive an implicit equation for the speed of invasion in terms of the Laplace transform of the CCTD. We obtain an expression for the minimum wave speed under this model and show that for a general CCTD the invasion can be arbitrarly fast.

In the second part of the paper, we focus our attention on MSMs. We study a spatially extended ABM which is designed to mimic cell invasion on a regular two-dimensional lattice. For each agent, we implement a general N-stage MSM to simulate the stochastic waiting time before the agent attempts to divide into two daughters. Through a mean-field closure approximation on the average agent density, we derive a system of N reaction-diffusion PDEs which represents a generalisation of the model of Vittadello et al. [2018]. By applying the front propagation method [Van Saarloos, 2003] to the system of PDEs, we reduce the computation of the invasion speed to an eigenvalue problem in terms of the rates of transition between consecutive stages, λ_i. We use this result to study the case of identical transition rates, that corresponds to modelling the CCTD as Erlang. In this case we provide the exact analytical expression for the speed. Finally, we formulate a result for the maximum and minimum speed for a general N-stage MSM.

The paper is organised as follows. In Section 2 we define the age-structured model and we derive the implicit equation for the invasion speed for general CCTD. In Section 3 we define two MSMs: a stochastic ABM and the corresponding mean-field approximation. In Section 3.1 we

7

explain how to apply the front propagation method and we state the eigenvalue problem. We present our results on Erlang distributed cell cycle times and the general hypoexponential case in Section 3.2. We conclude in Section 4 with a brief discussion of this work and future challenges.

2 Age-Structured Model

The Fisher-KPP equation implicitly assumes Markov dynamics for the individual cells making up the population, implying a cell cycle time with an exponential distribution [Fisher, 1937]. One way to adapt the model to allow for an arbitrary cell cycle time distribution is through the addition of age-structure. Cells have an associated age, denoted by a, which takes values in the positive real numbers and increases as time evolves. Cell motility is modelled as diffusion, with diffusivity D, and they proliferate with an age-dependent rate, $h(a)$.

We can write down a simple linear PDE for the density of cells with age a and spatial location x at time t, $C(a, x, t)$, as follows

$$
\frac{\partial}{\partial t} C(a, x, t) = -\frac{\partial}{\partial a} C(a, x, t) + D \frac{\partial^2}{\partial x^2} C(a, x, t) - h(a) C(a, x, t)
$$
$$
C(0, x, t) = 2 \int_0^\infty h(s) C(s, x, t) \mathrm{d}s \, .
\tag{2}
$$

The function $h(s)$ is the hazard rate, related to the probability density function $f(s)$ of the age at which cells divide (i.e. the CCTD) via

$$
h(s) = \frac{f(s)}{\int_s^\infty f(a)\, \mathrm{d}a} \, , \qquad f(s) = h(s) \exp\left(-\int_0^s h(a) \mathrm{d}a \right) \, .
\tag{3}
$$

The boundary condition for $C(0, x, t)$ in equation (2) gives the density of newborn cells as twice the total rate of cell division. Note that we have neglected from our formulation in system (2) any non-linear terms arising from crowding effects, as these are not relevant to the speed of the

8

front propagation. This model is a simple spatial adaptation of the McKendrick-Von Foerster equation for growing age-structured populations, and has been studied before [Webb and Webb, 1985, Al-Omari and Gourley, 2002, Gabriel et al., 2012].

As our first result, we show that the speed of propagation for the model (2) is determined by the Laplace transform of the CCTD, defined by

$$\mathcal{L}\{f\}(s) = \int_0^\infty e^{-sa} f(a) \, \mathrm{d}a \,. \tag{4}$$

Theorem 1 If $\lim_{s \to \infty} \mathcal{L}\{f\}(s) < 1/2$ then the PDE (2) admits travelling wave solutions with propagation speed $c > 2\sqrt{D\lambda}$, where $\lambda > 0$ is the unique solution to

$$\mathcal{L}\{f\}(\lambda) = 1/2 \,. \tag{5}$$

Proof. The system (2) is separable, hence we seek solutions of the form $C(a, x, t) = v(a)w(x - ct)$, corresponding to a travelling wave with speed c and internal age structure given by v. Inserting into (2) and rearranging, we find

$$c\frac{w'}{w} + D\frac{w''}{w} = \frac{v'}{v} + h \,. \tag{6}$$

The left-hand side here is a function only of $x - ct$, whilst the right-hand side is a function only of a. We thus determine that both are equal to a constant, say $-\lambda$. The w equation becomes

$$\lambda w + cw' + Dw'' = 0 \,, \tag{7}$$

which is well-known as the linearisation of the Fisher-KPP equation, admitting travelling wave

9

solutions for all $c > 2\sqrt{D\lambda}$. The equation for v has solution

$$v(a) = v(0) \exp\left(-a\lambda - \int_0^a h(\alpha)\mathrm{d}\alpha\right). \tag{8}$$

The boundary condition then gives us

$$1 = 2\int_0^\infty h(a) \exp\left(-a\lambda - \int_0^a h(\alpha)\mathrm{d}\alpha\right)\mathrm{d}a, \tag{9}$$

143 from which the definition of the hazard rate, equation (3), gives the result $1 = 2\mathcal{L}\{f\}(\lambda)$. Unique-
144 ness of the solution (when one exists) follows from the monotonicity of the Laplace transform of
145 a probability density. □

We can use the previous result to investigate the range of speeds for an arbitrary CCTD with a given mean, $\bar{\mu}$. By using Jensen's inequality we have that for any positive supported f with mean $\bar{\mu}$

$$\mathcal{L}\{f\}(\lambda) \leq e^{-\lambda\bar{\mu}} = \mathcal{L}\{\delta_{\bar{\mu}}\}(\lambda), \tag{10}$$

where $\delta_{\bar{\mu}}$ is the Dirac delta function concentrated at $\bar{\mu} > 0$. From the monotonicity of the Laplace transform of a probability density, it follows that the minimum speed is obtained by using $f = \delta_{\bar{\mu}}$, which gives

$$c \geq 2\sqrt{\frac{D\ln 2}{\bar{\mu}}} \tag{11}$$

We now use Theorem 1 to show that there is no upper bound for the speed of invasion of a general CCTD with a given mean. Consider the set of probability density functions defined as

$$f_\varepsilon(x) = \frac{1}{2}\left(\delta_{\varepsilon\bar{\mu}} + \delta_{(2-\varepsilon)\bar{\mu}}\right), \tag{12}$$

10

where $\varepsilon \leq 1$. It follows immediately that each member of this set of functions have mean $\bar{\mu}$ and Laplace transform given by:

$$\mathcal{L}\{f_\varepsilon\}(\lambda) = \frac{1}{2}\left(e^{-\lambda\varepsilon\bar{\mu}} + e^{-\lambda(2-\varepsilon)\bar{\mu}}\right).$$ (13)

By substituting the expression (13) into equation (5) and rearranging, we obtain the implicit equation for λ given by

$$\lambda\varepsilon\bar{\mu} = -\ln\left(1 - e^{-2\lambda\bar{\mu}}\right).$$ (14)

146 The right-hand side of equation (14) is a strictly decreasing function of λ that converges to 0 as
147 $\lambda \to \infty$. Therefore, we can always choose ε small enough so that the solution of equation (14) is
148 arbitrarily large.

This demonstrates that, assuming that the CCTD is a general function with mean $\bar{\mu}$ and positive support, the range of possible invasion speeds is given by

$$c \in \left[2\sqrt{\frac{D\ln 2}{\bar{\mu}}}, \infty\right).$$ (15)

149 The result in Theorem 1 is important because it establishes the connection between a general
150 CCTD and the corresponding invasion speed. However, for some particular classes of distribu-
151 tions, solving equation (5) analytically can be challenging and the method of this Section does
152 not provide any deeper insights. In particular, this is true for hypoexponential distributions,
153 which are of special interest in the context of cell proliferation. In remaining part of the paper
154 we further explore this class of distributions using a MSM of cell migration.

11

3   Multi-Stage Models

In this section we introduce the two MSM that we will use throughout the remainder of this paper. Firstly, we define a discrete ABM, in which the multi-stage representation of the CCTD is implemented as a stochastic feature of each cell at the microscale. Secondly, we introduce a system of deterministic PDEs describing the average cell density in a macroscopic manner.

The ABM   We consider a continuous-time ABM on a two-dimensional regular square lattice, with a given spacing denoted by Δ. Each cell is modelled as a single agent which moves and proliferates. Volume exclusion is incorporated by allowing at most one agent to occupy a given lattice site.

Agents move according to a simple excluding random walk on the lattice. Each agent attempts a movement after an exponentially distributed waiting time with rate α. When this happens, a new position is chosen uniformly from one of the four nearest neighbouring sites and the movement takes place only if the selected site is empty. The event is aborted otherwise.

We implement cell proliferation using a MSM. We divide the cell cycle into N sequential stages. Agents at one of the first $N-1$ stages, $i = 1, \ldots, N-1$, move to the next stage after an exponentially distributed waiting time of rate λ_i. Agents at the last stage, N, can attempt a proliferation event, after a further exponentially distributed waiting time of rate λ_N. In order to attempt a proliferation event, a target site is selected uniformly at random from one of the four nearest neighbouring sites. If such site is empty, a new first-stage agent is located on it, and the proliferating agent is returned to the first stage. If the target site is occupied, the proliferation event is aborted and the proliferating agent remains at the last stage[1].

---

[1]Alternatively, we could choose to return the proliferating agent to the first stage every time an abortion occurs. This model has been studied in Yates et al. [2017] for homogeneously distributed agents. This modification does not substantially change our results. This is because our analysis of the speed of the wave front is based on low density regions, where abortion of events does not play an important role. For this reason, we decided to focus only on the stated version of the model.

₁₇₆ We simulate the cell invasion by populating the first 10 columns of the lattice with agents at

₁₇₇ stages that are chosen uniformly at random. We impose zero flux boundary conditions on the

₁₇₈ x-direction and periodic boundary conditions on the y-direction. Agents are displaced uniformly

₁₇₉ at random in the vertical direction, so we can reduce the dimensionality of the problem by

₁₈₀ considering the average column density [Simpson et al., 2009].

₁₈₁ **The PDE model** Here we define the continuous model for the average column density which

₁₈₂ will be the object of the invasion speed analysis.

We denote by $S_i(x,t)$ the density of i-stage agents in the column x at time t, averaged over

multiple realisations of the ABM. Let $C(x,t)$ be the total density of column x at time t, i.e.

$$C(x,t) = \sum_{i=1}^{N} S_i(x,t). \tag{16}$$

By writing down the master equation of S_i, for $i = 1, \ldots, N$ and taking the limit as $\Delta \to 0$,

while keeping $\alpha\Delta^2$ constant, one can derive a system of reaction-diffusion PDEs for the column

densities of the different stages:

$$\begin{cases} \frac{\partial S_1}{\partial t} &= D\frac{\partial}{\partial x}\left[(1-C)\frac{\partial S_1}{\partial x} + S_1\frac{\partial C}{\partial x}\right] + 2\lambda_N(1-C)S_N - \lambda_1 S_1 \\[2mm] \frac{\partial S_i}{\partial t} &= D\frac{\partial}{\partial x}\left[(1-C)\frac{\partial S_i}{\partial x} + S_i\frac{\partial C}{\partial x}\right] + \lambda_{i-1}S_{i-1} - \lambda_i S_i \qquad \text{for } i = 2, \ldots, N-1 \\[2mm] \frac{\partial S_N}{\partial t} &= D\frac{\partial}{\partial x}\left[(1-C)\frac{\partial S_N}{\partial x} + S_N\frac{\partial C}{\partial x}\right] + \lambda_{N-1}S_{N-1} - \lambda_N(1-C)S_N \,, \end{cases} \tag{17}$$

₁₈₃ where $D = \lim_{\Delta\to 0}\frac{\alpha\Delta^2}{4}$. Notice that other types of tessellations than the regular square lattice

₁₈₄ are common in the literature [Deutsch and Dormann, 2007, Simpson et al., 2018]. However, the

₁₈₅ model formulation and the corresponding mathematical analysis in these cases do not change

₁₈₆ substantially. For example, a detailed derivation for the three-stage model on an hexagonal

13

187 lattice can be found in Simpson et al. [2018]. Moreover, we should underline that the diffusivity

188 of the cells in our model is independent of their stage which is not always true for real cells. We

189 discuss this and other possible generalisations of the model in Section 4.

System (17) consists of a set of N reaction-diffusion PDEs with non-linearities in both the diffusion and the proliferation terms due to the effect of volume exclusion. Specifically, the term $(1 - C)$ accounts for the reduction in rate due to volume exclusion. Notice that by summing all the equations in (17), we obtain

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} + \lambda_N(1 - C)S_N \,. \tag{18}$$

190 In other words, although the diffusion terms in each of the equations (17) are non-linear, the

191 motility at the population-level is simple diffusion [Simpson et al., 2009]. Conversely, due to the

192 dependence of equation (18) on S_N, it is not possible to obtain a closed PDE for the total agent

193 density without further assumptions.

194 We conclude this section by showing a comparison of the two models in Figure 2. In the

195 example, we choose realistic values of motility rate, α, and proliferation rate, $\bar{\lambda}$, as in Treloar

196 et al. [2013] and Haass et al. [2014], respectively. We consider an ABM with five stages with

197 increasing rates (the stage-to-stage transition rates are chosen to facilitate the visualisation of

198 the different density profiles). In panels (a), (b) and (c) three successive snapshots are shown

199 and the formation of the travelling wave appears clearly. As previously observed by Vittadello

200 et al. [2018], due to the presence of volume exclusion, the travelling wave solutions of the N

201 subpopulations of cells are of two qualitatively different types. The density profile of the first

202 $N - 1$ subpopulations have the form of moving pulses located at the front of the total wave with

203 the amplitude which depends on the rate of the corresponding stage. The profile of the last stage

204 subpopulation, instead, appears as a moving wavefront which dominates the density at the back
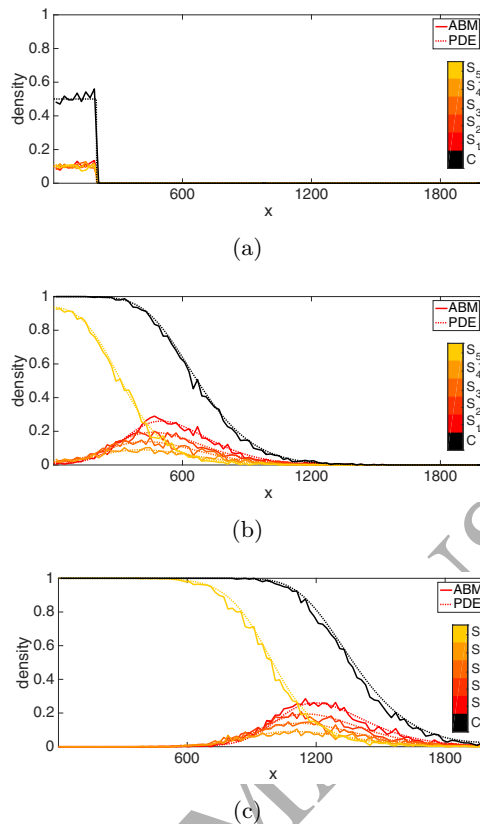
14

Figure 2: Comparison between the average column density for the ABM (full lines) and the PDE model (dotted lines) with a five-stage MSM. The panels show three snapshots of the evolution of the two models at time 0 (a), 150 (b) and 300 (c). In all cases, the profiles for the five different subpopulations are shown in different gradations of orange and the total density is plotted in black. The ABM profiles are obtained by averaging over 20 identically prepared simulations on a 2000×400 lattice. The other parameters of the models are $\Delta = 20$, $\alpha = 4$, $\bar{\lambda} = 0.0233$, $\lambda_1 = 0.15$, $\lambda_2 = 0.19$, $\lambda_3 = 0.25$, $\lambda_4 = 0.37$ and $\lambda_5 = 0.75$.

²⁰⁵ of the total wave.

²⁰⁶ The numerical solutions of the PDEs agree well with the average behaviour of the ABM.

²⁰⁷ Therefore, we focus our attention on the the speed of the PDE model which we can investigate

²⁰⁸ using an analytical approach (see Section 3.1).

²⁰⁹ It is important to note that the quantitative validity of our results on the PDE model will

²¹⁰ extend to the ABM only for the range of parameters which preserves the good agreement between
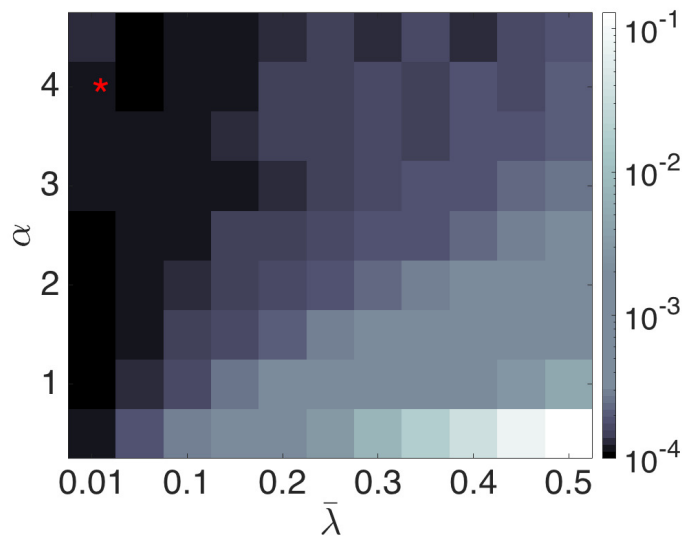
15

Figure 3: Histogram distance error (HDE) between column density profiles of ABM and the mean-field model for different parameters of movement and proliferation. The colour of each square denotes the HDE between the total density of the PDE and the ABM (averaging over 1000 simulations) as described in the legend. Dark squares represent small HDE which denotes good agreement between the two models, whereas light squares represent large HDE which denote a loss of agreement. The models are simulated on a 2000×400 domain with $\Delta = 20$ and the HDE is computed at the time when half of the domain in the mean-field model is occupied. All the ABMs are simulated with $N = 5$ and with transition rates proportional to the one of Figure 2. The red mark denotes the realistic parameter choice corresponding to Figure 2.

²¹¹ the two models. In Figure 3 we compare the total averaged column density profiles of the ABM

²¹² and PDE for different parameters. The heat map shows the histogram distance error (HDE)[2]

²¹³ between ABM and the PDE model for different rates of movement and proliferation. When

²¹⁴ the rate of proliferation is large compared to the motility rate, the mean-field approximation

²¹⁵ loses its accuracy. This is a well known phenomenon which is caused by the presence of strong

²¹⁶ spatial correlations between occupied sites, induced by the proliferation [Middleton et al., 2014].

²¹⁷ Increasing the motility parameter tends to break up spatial correlations of neighbouring sites

---

[2]The HDE between two normalised histograms with values a_i and b_i at point i (i.e. $\sum a_i = \sum b_i = 1$) is defined as $HDE = \sum |a_i - b_i|/2$ [Cao and Petzold, 2006].

16

218 and, consequentially, to improve the accuracy of the mean-field approximation.

219 The results of Figures 2 and 3 confirm that for realistic choice of parameters ($\alpha \approx 4$ and

220 $\bar{\lambda} \approx 0.02$ [Treloar et al., 2013, Haass et al., 2014]) the PDE model provides a good approximation

221 of the ABM. This motivates us to focus our analysis on the continuum model. It is possible to

222 derive more accurate descriptions in those cases where the agreement is lost using higher order

223 moment closure schemes (see for example Baker and Simpson [2010], Markham et al. [2013]),

224 but this is beyond the scope of this paper.

225 **3.1 Wavespeed Analysis**

226 In this Section we apply the front propagation method of Van Saarloos [2003] to system (17) to

227 study the speed of invasion of the PDE model.

228

The system of equations (17) has two equilibria, an unstable empty state, $S_i(x,t) \equiv 0$ for

$i = 1, \ldots, N$, and a stable occupied state, $S_i(x,t) \equiv 0$ for $i = 1, \ldots, N-1$ and $S_N(x,t) \equiv 1$.

Firstly we linearise the system about the unstable steady state, giving

$$\begin{cases} \frac{\partial S_1}{\partial t} = D\frac{\partial^2 S_1}{\partial x^2} + 2\lambda_N S_N - \lambda_1 S_1 \\ \\ \frac{\partial S_i}{\partial t} = D\frac{\partial^2 S_i}{\partial x^2} + \lambda_{i-1}S_{i-1} - \lambda_i S_i \qquad \text{for } i = 2, \ldots, N \ . \end{cases} \tag{19}$$

We substitute

$$S_i(x,t) \propto \exp\left(-\iota\omega(k)t + \iota kx\right) ,$$

into equations (19), where ι is the immaginary unit, $\omega(k)$ is the dispersion angular frequency of

17

the Fourier modes and k is the spatial wavenumber. Upon simplification, we obtain

$$
\begin{cases}
-\iota\omega(k) & = -Dk^2 + 2\lambda_N - \lambda_1 \\[2mm]
-\iota\omega(k) & = -Dk^2 + \lambda_{i-1} - \lambda_i \qquad \text{for } i = 2, \ldots, N.
\end{cases}
$$

Following the front propagation method [Van Saarloos, 2003], the expression of the wave speed, c, is given by

$$
c = \frac{\text{Im}\left[\omega(k^*)\right]}{\text{Im}\left[k^*\right]}, \tag{20}
$$

where $k^* = \iota q$, with q real, and such that

$$
\frac{\mathrm{d}\omega}{\mathrm{d}k}(k^*) = \frac{\text{Im}\left[\omega(k^*)\right]}{\text{Im}\left[k^*\right]}. \tag{21}
$$

Notice that we can write down $\iota\omega(k)$ in the form

$$
\iota\omega(k) = k^2 D - \rho, \tag{22}
$$

where ρ is an eigenvalue of the matrix

$$
\Lambda = \begin{bmatrix}
-\lambda_1 & 0 & \ldots & 0 & 2\lambda_N \\
\lambda_1 & -\lambda_2 & 0 & \ldots & 0 \\
0 & \lambda_2 & -\lambda_3 & \ldots & 0 \\
\vdots & & \ddots & \ddots & \vdots \\
0 & \ldots & & \lambda_{N-1} & -\lambda_N
\end{bmatrix}. \tag{23}
$$

18

From expression (22) it follows that

$$\frac{\mathrm{d}\omega}{\mathrm{d}k}(k^*) = 2qD \,, \tag{24a}$$

$$\frac{\mathrm{Im}\,[\omega(k^*)]}{\mathrm{Im}\,[k^*]} = \frac{q^2 D + \mathrm{Re}\,[\rho]}{q} \,. \tag{24b}$$

By substituting equations (24) into (21), we obtain $q^2 = \mathrm{Re}\,[\rho]/D$. Note that Λ is Metzler, and hence by Perron-Fobenius its rightmost eigenvalue is real. Hence, from (20), the wave speed of the invasion is given by

$$c = 2\sqrt{D\rho} \,, \tag{25}$$

where ρ is the maximum real eigenvalue of Λ, defined in terms of the characteristic polynomial of the matrix Λ, $\mathcal{P}_\Lambda(x)$, as follows

$$\rho(\Lambda) = \max\left\{x \in \mathbb{R} \mid \mathcal{P}_\Lambda(x) = 0\right\} \,. \tag{26}$$

229 This shows that the problem of finding the speed of invasion of the PDE model is equivalent to
230 computing the maximum real eigenvalue of the matrix Λ, $\rho(\Lambda)$.

231 **3.2 Results**

The characteristic polynomial of the matrix Λ can be computed directly from the matrix and it reads

$$\mathcal{P}_\Lambda(x) = \prod_{i=1}^{N} (\lambda_i + x) - 2\prod_{i=1}^{N} \lambda_i \,. \tag{27}$$

232 In general, an analytical formula of the roots of the polynomial function $\mathcal{P}_\Lambda(x)$ is not available.
233 In this section we first consider the case of $\lambda_i = \lambda$ for $i = 1, \ldots, N$ for which the maximum
234 eigenvalue $\rho(\Lambda)$ can be computed analytically. This corresponds to a special case of the general

19

235 hypoexponential distribution, known as the Erlang distribution. We conclude by proving a

236 theorem in which we state the range of speed variability for the general hypoexponential CCTD.

The Erlang distribution Consider the case $\lambda_i = \lambda$ for $i = 1, \ldots, N$, which corresponds the Erlang CCTD. Under this assumption, we can write down the characteristic equation of the matrix Λ, using formula (27), as

$$(\lambda + x)^N = 2\lambda^N. \tag{28}$$

The eigenvalues of Λ are then given by the solutions of equation (28) which are $x_j = \lambda \left(\xi^j \sqrt[N]{2} - 1 \right)$ for $j = 1, \ldots, N$, where $\xi = \exp\left(2\pi\iota/N\right)$ is the primitive N-th root of unity. Hence, we obtain that

$$\rho(\Lambda) = \lambda \left(\sqrt[N]{2} - 1 \right). \tag{29}$$

By substituting the expression (29) into equation (25) we obtain the formula for the speed of invasion for the model with Erlang distribution

$$c = 2\sqrt{D\lambda \left(\sqrt[N]{2} - 1 \right)}. \tag{30}$$

237 Notice that for $N = 1$, which corresponds to exponential CCTD, we recover the well known

238 expression of the speed for the Fisher-KPP equation, $2\sqrt{D\lambda}$.

239 **The general case** For the case of a general hypoexponential distribution, there is no analytical

240 formula for the expression of the maximum real eigenvalue of the matrix Λ. However, we find

241 that the Erlang case and the exponential case, for which we do have the analytical formula of

242 the speed, correspond to the lower and upper bound (respectively) for the speed of travelling

243 waves with hypoexponential CCTD and a given total proliferation rate, $\bar{\lambda}$. This result follows

20

244 directly from the following theorem on the range of $\rho(\Lambda)$.

Theorem 2 Let $\rho(\Lambda)$ be defined by equation (26) as the maximum real eigenvalue of the matrix Λ. Then

$$\bar{\lambda} N \left(\sqrt[N]{2} - 1 \right) \leq \rho(\Lambda) < \bar{\lambda} \,, \tag{31}$$

245 where $\bar{\lambda} = \left(\sum_{i=1}^{N} 1/\lambda_i \right)^{-1}$.

A proof of Theorem 2 can be found in the appendix. It is immediate to interpret the result of the Theorem 2 in terms of invasion speeds. In particular, by using equation (25), together with the two inequalities (31), we deduce that the speed of the invasion of the PDE model with diffusion coefficient D and a general N-stage representation of the CCTD with total growth rate given by $\bar{\lambda}$, lies in the interval

$$c \in \left[2\sqrt{D\bar{\lambda} N \left(\sqrt[N]{2} - 1 \right)}, \, 2\sqrt{D\bar{\lambda}} \right) \,. \tag{32}$$

We can generalise this result even further by taking the limit as $N \to \infty$ in the right-hand side of equation (32). Hence we obtain a general interval which holds for any multi-stage representation, regardless of the number of stages, which reads

$$c \in \left(2\sqrt{D\bar{\lambda} \ln 2}, \, 2\sqrt{D\bar{\lambda}} \right) \,, \tag{33}$$

246 where we used $N \left(\sqrt[N]{2} - 1 \right) = \ln 2 + \mathcal{O} \left(N^{-1} \right)$.

247 Notice that the lower bound of the interval (33) is equivalent to the lower bound for the
248 general CCTD, obtained in (15) of Section 2. This can be intuitively understood by observing
249 that, as we let number of stages of an hypoexponential distribution go to infinity while keeping the
250 total rate, $\bar{\lambda}$, fixed, the variance of the distribution tends to zero. Consequently, the distribution
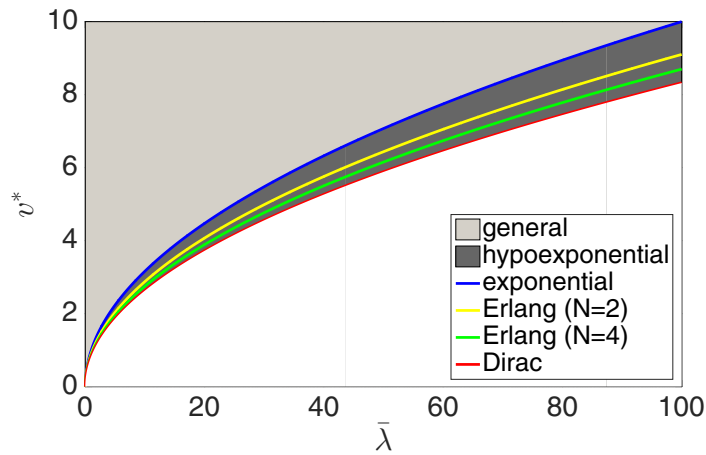
21

Figure 4: Illustration of the range of invasion speeds for a fixed mean proliferation rate and diffusion coefficient, $D = 1$. The two coloured regions represent the range of speed for a general CCTD. The dark grey subregion highlights the range of speeds for hypoexponential CCTDs. The global minimum speed is obtained by using the Dirac distribution (red line). The exponential CCTD (blue line) is the hypoexponential distribution which leads to maximum speed. There is no upper bound for the general case. Two examples of Erlang CCTDs with two stages (yellow line) and four stages (green line) are also shown.

converges to a Dirac function concentrated in the mean, $\bar{\mu} = \bar{\lambda}^{-1}$, which we have proved in Section 2 to be the distribution corresponding to the minimum invasion speed. In Figure 4 we summarise our findings about the range of invasion speed for different CCTD through a graphical representation.

4   Conclusion

In this work we investigated the quantitative effect of implementing a realistic CCTD into models of cell invasion. Firstly, we derived a general result from a generalised version of the Fisher-KPP equation. Then we investigated the case of MSMs by implementing a simple ABM of cells undergoing undirected migration and proliferation by division, in which the time between successive divisions is modelled using a multi-stage representation (i.e. the CCTD is hypoexponential). By

22

₂₆₁ studying a continuous version of the ABM, we connected the type of CCTD to the speed of the

₂₆₂ corresponding invasion.

₂₆₃ The results indicate that, for a fixed mean division time, the minimum speed of invasion is

₂₆₄ obtained by the Dirac distribution, while there is no upper bound. In other words, the invasion

₂₆₅ can be, in general, infinitely fast. However, when we focus our attention to the case of MSMs,

₂₆₆ which are known to represent well the experimental CCTD, our analysis shows that the speed can

₂₆₇ vary in a bounded interval (see Figure 4). More precisely, we show that the maximum invasion

₂₆₈ speed is reached by adopting an exponential CCTD, which leads to the classic the Fisher-KPP

₂₆₉ model. On the other hand, the minimum speed is obtained by partitioning the CCTD into

₂₇₀ multiple exponential stages with identical rates, which corresponds to the case of Erlang CCTD.

₂₇₁ Finally, by considering the limiting case of infinitely many stages, we find that the infimum value

₂₇₂ of the speed for the class of hypoexponential CCTD coincides with the global minimum for a

₂₇₃ general CCTD.

₂₇₄ The results indicate the invasion speed changes with the variance of the CCTD, i.e. decreasing

₂₇₅ the variance in the proliferation time distribution leads to slower invasion. We found that the

₂₇₆ maximum reduction in comparison to the classical formula for the Fisher-KPP model, is given

₂₇₇ by a multiplicative factor of $\sqrt{\ln 2} \approx 0.83$. Whilst interpreting this result in the context of

₂₇₈ experimental data is beyond the aim of this work, we want to stress that for number of stages

₂₇₉ $N \gg 1$, which is typically the case for experimentally observed distributions [Golubev, 2016,

₂₈₀ Yates et al., 2017, Chao et al., 2018], the speed converges to the lower bound of equation (33)

₂₈₁ with order given by $\mathcal{O}\left(N^{-1}\right)$. This suggests that, with the only information of the mean of the

₂₈₂ CCTD (equivalently, the total rate), including the factor $\sqrt{\ln 2}$ in the formula for the speed leads

₂₈₃ to a more accurate estimation than the classic expression of Fisher-KPP.

₂₈₄ In Section 3 we used a discrete ABM, but it is important to notice that alternative modelling

23

₂₈₅ approaches might lead to different results. Although discrete space ABMs are widespread in

₂₈₆ the literature [Deutsch and Dormann, 2007, Simpson et al., 2007, Cheeseman et al., 2014, Mort

₂₈₇ et al., 2016, Simpson et al., 2018, Vittadello et al., 2018], a considerable number of studies focus

₂₈₈ on lattice-free ABMs in which cells' positions are not constrained to a grid [Grima, 2008, Dyson

₂₈₉ et al., 2012, Dyson and Baker, 2014, Middleton et al., 2014, Matsiaka et al., 2017]. Another

₂₉₀ alternative approach, known as compartment-based model, consists in allowing multiple cells

₂₉₁ occupying a single lattice site [Taylor et al., 2015, 2016, Cianci et al., 2017]. In the context of our

₂₉₂ work, adopting alternative modelling regimes, such as lattice-free or compartment-based models,

₂₉₃ would lead to different nonlinear factors in system (17). Since the analysis of the wave speed is

₂₉₄ based on a linearisation of system (17), we believe that our results would still hold qualitatively.

₂₉₅ However, a rigorous comparison of these modelling approaches is beyond the scope of this paper.

₂₉₆ An important question that remains unanswered is the role of motility heterogeneity within

₂₉₇ the cell cycle. Experimental studies have found that the motility of a cell can depend on its

₂₉₈ cell cycle phase [Vittadello et al., 2018]. For example, during the mitotic phase, cells tend to

₂₉₉ reduce their movement [Mort et al., 2016]. In order to investigate this phenomenon in the light

₃₀₀ of the invasion speed, we could modify our model to allow different diffusion coefficients, D_i for

₃₀₁ $i = 1, \ldots, N$, for each stage in the system (17). Another aspect of the cell movement that can

₃₀₂ vary within the cell cycle is the directional persistence. Our models do not incorporate directional

₃₀₃ persistence of cells. However, it is possible to combine a MSM with existing models of directional

₃₀₄ persistence [Codling et al., 2008, Gavagnin and Yates, 2018]. Unfortunately, the application of

₃₀₅ the front propagation method of Van Saarloos [2003] (see Section 3.1) to these models leads to

₃₀₆ a dead end and it may be necessary to study the problem using a different approach. We will

₃₀₇ investigate this in future research.

24

Acknowledgments

References

J. Al-Omari and S.A. Gourley. Monotone travelling fronts in an age-structured reaction-diffusion model of a single species. *J. Math. Biol.*, 45(4):294–312, 2002.

A.R.A. Anderson and M.A.J. Chaplain. Continuous and discrete mathematical models of tumor-induced angiogenesis. *Bull. Math. Biol.*, 60(5):857–899, 1998.

R.E. Baker and M.J. Simpson. Correcting mean-field approximations for birth-death-movement processes. *Phys. Rev. E*, 82(4):041905, 2010.

Y. Cao and L. Petzold. Accuracy limitations and the measurement of errors in the stochastic simulation of chemically reacting systems. *J. Comp. Phys.*, 212(1):6–24, 2006.

H.X. Chao, R.I. Fakhreddin, H.K. Shimerov, R.J. Kumar, G.P. Gupta, and J.E. Purvis. Evidence that the cell cycle is a series of uncoupled, memoryless phases. *bioRxiv*, page 283614, 2018.

B.L. Cheeseman, D.F. Newgreen, and K.A. Landman. Spatial and temporal dynamics of cell generations within an invasion wave: A link to cell lineage tracing. *J. Theor. Biol.*, 363: 344–356, 2014.

C. Cianci, S. Smith, and R. Grima. Capturing brownian dynamics with an on-lattice model of hard-sphere diffusion. *Phys. Rev. E*, 95(5):052118, 2017.

E.A. Codling, M.J. Plank, and S. Benhamou. Random walk models in biology. *J. R. Soc. Interface*, 5(25):813–834, 2008.

25

329 A. Csikász-Nagy, D. Battogtokh, K.C. Chen, B. Novák, and J.J. Tyson. Analysis of a generic
330   model of eukaryotic cell-cycle regulation. *Biophys. J.*, 90(12):4361–4379, 2006.

331 M. Deng, W.L. Chen, A. Takatori, Z. Peng, L. Zhang, M. Mongan, R. Parthasarathy, M. Sartor,
332   M. Miller, J. Yang, et al. A role for the mitogen-activated protein kinase kinase kinase 1 in
333   epithelial wound healing. *Mol. Biol. Cell*, 17(8):3446–3455, 2006.

334 A. Deutsch and S. Dormann. *Cellular automaton modeling of biological pattern formation: char-*
335   *acterization, applications, and analysis*. Springer Science & Business Media, 2007.

336 L. Dyson and R.E. Baker. The importance of volume exclusion in modelling cellular migration.
337   *J. Math. Biol.*, 2014.

338 L. Dyson, P.K. Maini, and R.E. Baker. Macroscopic limits of individual-based models for motile
339   cell populations with volume exclusion. *Phys. Rev. E*, 86(3):031903, 2012.

340 E.C. Elliott and S.J. Cornell. Dispersal polymorphism and the speed of biological invasions.
341   *PLoS one*, 7(7):e40496, 2012.

342 R.A. Fisher. The wave of advance of advantageous genes. *Ann. Hum. Genet.*, 7(4):355–369,
343   1937.

344 P. Gabriel, S.P. Garbett, V. Quaranta, D.R. Tyson, and G.F. Webb. The contribution of age
345   structure to cell population responses to targeted therapeutics. *J. Theor. Biol.*, 311:19–27,
346   2012.

347 E. Gavagnin and C.A. Yates. Modeling persistence of motion in a crowded environment: The
348   diffusive limit of excluding velocity-jump processes. *Phys. Rev. E*, 97:032416, 2018.

349 C. Gérard and A. Goldbeter. Temporal self-organization of the cyclin/cdk network driving the
350   mammalian cell cycle. *Proc. Natl. Acad. Sci.*, 106(51):21643–21648, 2009.

26

S.F. Gilbert. The morphogenesis of evolutionary developmental biology. *Int. J. Dev. Biol.*, 47 (7-8):467, 2003.

A. Golubev. Applications and implications of the exponentially modified gamma distribution as a model for time variabilities related to cell proliferation and gene expression. *J. Theor. Biol.*, 393:203–217, 2016.

V. Gray-Schopfer, C. Wellbrock, and R. Marais. Melanoma biology and new targeted therapy. *Nature*, 445(7130):851, 2007.

R. Grima. Multiscale modeling of biological pattern formation. *Curr. Top. Dev. Biol*, 81:435–460, 2008.

N.K. Haass and B. Gabrielli. Cell cycle-tailored targeting of metastatic melanoma: Challenges and opportunities. *Exp. Dermatol.*, 26(7):649–655, 2017.

N.K. Haass, K.A. Beaumont, D.S. Hill, A. Anfosso, P. Mrass, M.A. Munoz, I. Kinjyo, and W. Weninger. Real-time cell cycle imaging during melanoma growth, invasion, and drug response. *Pigment Cell Melanoma Res.*, 27(5):764–776, 2014.

D. Hanahan and R.A. Weinberg. The hallmarks of cancer. *Cell*, 100(1):57–70, 2000.

R. Keller. Cell migration during gastrulation. *Curr. Opin. Cell Biol.*, 17(5):533–541, 2005.

Piscounov N. Kolmogorov A., Petrovskii I. A study of the diffusion equation with increase in the amount of substance, and its application to a biological problem. In V.M. Tikhomirov, editor, *Selected works of AN Kolmogorov*, pages 248–270. Springer, 1991.

P.K. Maini, D.L. S. McElwain, and D.I. Leavesley. Traveling wave model to interpret a wound-healing cell migration assay for human peritoneal mesothelial cells. *Tissue Eng.*, 10(3-4): 475–482, 2004.

27

373 D.C. Markham, M.J. Simpson, P.K. Maini, E.A. Gaffney, and R.E. Baker. Incorporating spatial

374 correlations into multispecies mean-field models. *Phys. Rev. E*, 88(5):052713, 2013.

375 O.M. Matsiaka, C.J. Penington, R.E. Baker, and M.J. Simpson. Continuum approximations for

376 lattice-free multi-species models of collective cell migration. *J. Theor. Biol.*, 422:1–11, 2017.

377 A.M. Middleton, C. Fleck, and R. Grima. A continuum approximation to an off-lattice individual-

378 cell based model of cell migration and adhesion. *J. Theor. Biol.*, 359:220–232, 2014.

379 R.L. Mort, M.J. Ford, A. Sakaue-Sawano, N.O. Lindstrom, A. Casadio, A.T. Douglas, M.A.

380 Keighren, P. Hohenstein, A. Miyawaki, and I.J. Jackson. Fucci2a: a bicistronic cell cycle

381 reporter that allows Cre mediated tissue specific expression in mice. *Cell Cycle*, 13(17):2681–

382 2696, 2014.

383 R.L. Mort, R.J.H. Ross, K.J. Hainey, O.J. Harrison, M.A. Keighren, G. Landini, R.E. Baker, K.J.

384 Painter, I.J. Jackson, and C.A. Yates. Reconciling diverse mammalian pigmentation patterns

385 with a fundamental mathematical model. *Nat. Commun.*, 7, 2016.

386 J.D. Murray. *Mathematical biology: I. An introduction*, volume 17. Springer Science & Business

387 Media, 2007.

388 M.G. Neubert and H. Caswell. Demography and dispersal: calculation and sensitivity analysis

389 of invasion speed for structured populations. *Ecology*, 81(6):1613–1628, 2000.

390 G.G. Powathil, K.E. Gordon, L.A. Hill, and M.A.J. Chaplain. Modelling the effects of cell-cycle

391 heterogeneity on the response of a solid tumour to chemotherapy: biological insights from a

392 hybrid multiscale cellular automaton model. *J. Theor. Biol.*, 308:1–19, 2012.

393 H.M. Sadeghi, B. Seitz, S. Hayashi, L. LaBree, and P.J. McDonnell. In vitro effects of mitomycin-

394 c on human keratocytes. *J. Refract. Surg.*, 14(5):534–540, 1998.

28

395 M.J. Simpson, A. Merrifield, K.A. Landman, and B.D. Hughes. Simulating invasion with cellular

396 automata: connecting cell-scale and population-scale properties. *Phys. Rev. E*, 76(2):021918,

397 2007.

398 M.J. Simpson, K.A. Landman, and B.D. Hughes. Multi-species simple exclusion processes. *Phys.*

399 *A*, 388(4):399–406, 2009.

400 M.J. Simpson, W. Jin, S.T. Vittadello, T. Tambyah, J. Ryan, G. Gunasingh, N. Haass, and

401 S. McCue. Stochastic models of cell invasion with fluorescent cell cycle indicators. *Phys. A*,

402 510:375–386, 2018.

403 P.R. Taylor, C.A. Yates, M.J. Simpson, and R.E. Baker. Reconciling transport models across

404 scales: the role of volume exclusion. *Phys. Rev. E*, 92(040701), 2015.

405 P.R. Taylor, R.E. Baker, M.J. Simpson, and C.A. Yates. Coupling volume-excluding

406 compartment-based models of diffusion at different scales: Voronoi and pseudo-compartment

407 approaches. *J. R. Soc. Interface*, 13(120), 2016. ISSN 1742-5689.

408 K.K. Treloar, M.J. Simpson, P. Haridas, K.J. Manton, D.I. Leavesley, D.S. McElwain, and R.E.

409 Baker. Multiple types of data are required to identify the mechanisms influencing the spatial

410 expansion of melanoma cell colonies. *BMC Syst. Biol.*, 7(1):137, 2013.

411 W. Van Saarloos. Front propagation into unstable states. *Phys. Rep.*, 386(2-6):29–222, 2003.

412 S.T Vittadello, S.W. McCue, G. Gunasingh, N.K Haass, and M.J. Simpson. Mathematical

413 models for cell migration with real-time cell cycle dynamics. *Biophys. J.*, 114(5), 2018.

414 G.F. Webb and G. Webb. *Theory of nonlinear age-dependent population dynamics*. CRC Press,

415 1985.

416 S.M. Wise, J.S. Lowengrub, H.B. Frieboes, and V. Cristini. Three-dimensional multispecies

417 nonlinear tumor growth–I: model and numerical method. *J. Theor. Biol.*, 253(3):524–543,

418 2008.

419 C.A. Yates, M.J. Ford, and R.L. Mort. A multi-stage representation of cell proliferation as a

420 markov process. *Bull. Math. Biol.*, 79(12):2905–2928, 2017.

421   Appendix

A   Proof of Theorem 2

Proof. Let $\mu_i = 1/\lambda_i$ for every $i = 1, \ldots, N$. By writing the characteristic equation $\mathcal{P}_\Lambda(x) = 0$ in terms of the parameters μ_i and upon rearranging, we obtain

$$\prod_{i=1}^{N} (\mu_i x + 1) = 2 \,. \tag{A.1}$$

We can write $\rho(\Lambda) = \rho(\mu_1, \ldots, \mu_N) = \rho(\underline{\mu})$ as

$$\rho(\underline{\mu}) = \max \left\{ x \in \mathbb{R} \mid \prod_{i=1}^{N} (\mu_i x + 1) = 2 \right\} , \tag{A.2}$$

423   for every $\underline{\mu} \in \{\mathbb{R}_{>0}\}^N$. It is easy to observe that $\rho(\underline{\mu})$ is a positive continuous function and we

424   can extend the definition (A.2) to $\underline{\mu} \in \{\mathbb{R}_{\geq 0}\}^N \setminus \{(0, \ldots, 0)\}$, by continuity.

Now fix $\bar{\lambda} = \left(\sum_{i=1}^{N} \mu_i \right)^{-1}$; without loss of generality we can take $\sum_{i=1}^{N} \mu_i = 1$, whence (31) becomes $N \left(\sqrt[N]{2} - 1 \right) \leq \rho(\underline{\mu}) < 1$. The case of general $\bar{\lambda}$ follows by multiplying by rescaling factor. Since ρ is a continuous function, we aim to find the stationary points of $\rho(\underline{\mu})$ in the N-dimensional simplex:

$$\mathcal{U}_N = \left\{ (\mu_1, \ldots, \mu_N) \in (0, 1]^N \; \middle| \; \sum_{i=1}^{N} \mu_i = 1 \right\} . \tag{A.3}$$

We apply the Lagrange multipliers method. Hence we study the Lagrangian function given by

$$\mathcal{L}(\mu_1, \ldots, \mu_N, \sigma) = \rho(\underline{\mu}) + \sigma \left(\sum_{i=1}^{N} \mu_i - 1 \right) . \tag{A.4}$$

31

Throughout we adopt the notation $\mathcal{L}_j = \frac{\partial \mathcal{L}}{\partial \mu_j}$ and $\rho_j = \frac{\partial \rho}{\partial \mu_j}$. By imposing $\mathcal{L}_j = 0$ we obtain

$$\rho_j = -\sigma \,, \tag{A.5}$$

for all $j = 1, \ldots, N$. We can now differentiate equation (A.1) respect to μ_j, which gives us

$$0 = \sum_{i=1}^{N} \prod_{k \neq i} (1 + \mu_k \rho) \, (\rho \, \delta_{i,j} + \mu_i \rho_j) \tag{A.6}$$

where $\delta_{i,j}$ denotes the Kronecker delta. If we multiply and divide each term of the right-hand side of equation (A.6) by $(1 + \mu_i \rho)$, we obtain

$$0 = \sum_{i=1}^{N} \frac{\rho \, \delta_{i,j} + \mu_i \rho_j}{1 + \mu_i \rho}$$
$$= \frac{\rho}{1 + \mu_j \rho} + \rho_j \sum_{i=1}^{N} \frac{\mu_i}{1 + \mu_i \rho} \,, \tag{A.7}$$

By combining equations (A.5) and (A.7) we gain a condition on the coordinate μ_j of the stationary points, namely

$$\frac{\rho}{1 + \mu_j \rho} = \sigma \sum_{i=1}^{N} \frac{\mu_i}{1 + \mu_i \rho} \,. \tag{A.8}$$

425 Notice that equation (A.8) holds for every $j = 1, \ldots, N$ and it is independent of j, hence the
426 only stationary point of $\rho(\underline{\mu})$ in the simplex $\mathcal{U}_N$ is the given by the centre $\underline{\mu}_N^* = (1/N, \ldots, 1/N)$.

To conclude we need study the value of $\rho(\underline{\mu})$ on the boundary of the simplex, defined as

$$\partial \mathcal{U}_N = \left\{ (\mu_1, \ldots, \mu_N) \in [0,1]^N \, \Big| \, \sum_{i=1}^{N} \mu_i = 1 \text{ and } \mu_j = 0, \, \exists j \in \{1, \ldots, N\} \right\} . \tag{A.9}$$

Let us consider the elements of $\partial \mathcal{U}_N$ with exactly n non-zero coordinates, with $n = 1, \ldots, N-1$.

32

Without loss of generality we can focus on the points of the form

$$(\mu_1, \ldots, \mu_n, 0, \ldots, 0) \in \partial \mathcal{U}_N \,, \tag{A.10}$$

where $(\mu_1, \ldots, \mu_n) \in \mathcal{U}_n$. Notice that the $\rho(\underline{\mu})$ is well defined in such points by continuity, as observed before. By repeating the Lagrange multiplier method in the sub-simplex $\mathcal{U}_n$, we find that the only stationary point of $\rho(\underline{\mu})$ of the form (A.10) is the one with $\mu_1 = \mu_2 = \cdots = \mu_n$, i.e.:

$$\underline{\mu}_n^* = (\underbrace{1/n, \ldots, 1/n}_{n}, 0, \ldots, 0) \in \partial \mathcal{U}_N. \tag{A.11}$$

427 This holds for every $n = 1, \ldots, N-1$, so we can write all the stationary points of $\rho(\underline{\mu})$ in $\partial \mathcal{U}_N$
428 upon permutation of the coordinates in the form (A.11).

All the stationary points $\underline{\mu}_n^*$, for $n = 1, \ldots N$, correspond to an Erlang distribution for which we can compute the expression of ρ directly from the definition (A.2) as

$$\rho(\underline{\mu}_n^*) = n \left(\sqrt[n]{2} - 1 \right) , \tag{A.12}$$

for $n = 1, \ldots, N$. The right-hand side of equation (A.12) is a decreasing function of n. We deduce that the centre of the simplex, $\mu_N^* \in \mathcal{U}_N$, corresponds to the global minimum, i.e. for all $\underline{\mu} \in \mathcal{U}_N$

$$\rho(\underline{\mu}) \geq \rho(\underline{\mu}_N^*) = N \left(\sqrt[N]{2} - 1 \right) . \tag{A.13}$$

Finally, $\mu_1^* \in \partial \mathcal{U}_N$ and all the points obtained by permuting its coordinates, correspond to

33

supremum points, i.e. for all $\underline{\mu} \in \mathcal{U}_N$

$$\rho(\underline{\mu}) < \rho(\underline{\mu}_1^*) = 1 \,. \tag{A.14}$$

429 $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$