bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Stochastic simulation tools and continuum models for describing two-dimensional collective cell spreading with universal growth functions

Wang Jin, Catherine J. Penington, Scott W. McCue, and Matthew J. Simpson*.

School of Mathematical Sciences, Queensland University of Technology (QUT), Brisbane, Australia.

Abstract

Two-dimensional collective cell migration assays are used to study malignant spreading and tissue repair. These assays involve combined cell migration and cell proliferation processes, both of which are modulated by cell-to-cell crowding effects. Previous discrete models of two-dimensional collective cell migration assays involve a nearest-neighbour proliferation mechanism where crowding effects are incorporated by aborting potential proliferation events if the randomly chosen target site is occupied. There are two potential limitations of this traditional approach: (i) it seems unreasonable to abort a potential proliferation event based on the occupancy of a single, randomly chosen target site; and, (ii) the continuum limit description of this mechanism leads to the standard logistic growth function, but there is increasing evidence suggesting that cells do not always grow logistically. Motivated by these limitations we introduce a generalised proliferation mechanism into a two-dimensional lattice-based exclusion process model, which allows non-nearest neighbour proliferation events to take place over a template of $r \geq 1$ concentric rings of lattice sites. Further, the decision to abort potential proliferation events is made using a *crowding function*, $f(C) \in [0, 1]$ with $f(0) = 1$ and $f(1) = 0$. This approach accounts for the density of agents within a group of lattice sites rather than dealing with the occupancy of a single site only. Analysing the continuum limit description of the stochastic model shows that the standard logistic source term, $\lambda C(1 - C)$, where λ is the proliferation rate, is generalised to a universal growth function, $\lambda C f(C)$. Comparing the solution of the continuum description with averaged simulation data indicates that the continuum model performs well for many choices of $f(C)$ and r. For nonlinear $f(C)$, the quality of the continuum-discrete match increases with r. Therefore, we suggest that estimating r from time lapse images will help distinguish between situations where the simpler continuum model is adequate from other situations where repeated simulations of the stochastic algorithm is required.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

I. INTRODUCTION

Two-dimensional collective cell migration assays are routinely used to study combined cell migration and cell proliferation processes [1, 2]. These assays provide insight into cancer [3, 4] and tissue repair [5, 6]. Broadly speaking, there are two different kinds of two-dimensional collective cell migration assays. (i) *Cell proliferation assays*, as shown in Figure 1(a)-(c), are initiated by uniformly distributing a population of cells on a two-dimensional substrate. Over time, individual cells undergo migration and proliferation events, which eventually leads to the formation of a confluent cell monolayer [7, 8]. (ii) *Scratch assays*, as shown in Figure 1(d)-(f), are also initiated by uniformly distributing a population of cells on a two-dimensional substrate. However, a wound, or scratch is made to create a cell-free region adjacent to a region that is occupied by cells [9, 10]. Over time, individual cells undergo motility and proliferation events with the net result being the spreading of cells into the vacant region [9, 10]. A critical feature of both proliferation and scratch assays is the role of cell-to-cell interactions and crowding. At low cell density, individual cells are relatively free to move and proliferate because of the abundance of free space [7, 12]. In contrast, at high cell density, individual cells are strongly influenced by cell-to-cell crowding, which reduces their ability to move and proliferate [7, 12].

There are two different approaches to modelling two-dimensional collective cell migration assays. In the first approach, a continuum reaction-diffusion equation is applied to mimic certain features of the experiment [5, 6]. Using continuum models, most previous studies represent cell migration with a diffusion-type mechanism and a logistic source term to represent carrying capacity-limited proliferation [5, 6, 12–19]. In the second modelling approach, a discrete random walk model is used to mimic certain features of these experiments [20, 21]. Here, many previous studies represent cell migration using an unbiased exclusion process [22], which incorporates hard core exclusion to model cell-to-cell crowding [23–28]. Cell proliferation is incorporated by allowing agents to place daughter agents on the lattice, and crowding effects can be incorporated by ensuring that potential proliferation events that would place a daughter agent on an occupied site are aborted [25, 26]. Discrete random walk models have an advantage over continuum models when it comes to comparing model predictions with experimental observations. Experimental data in the form of snapshots and time lapse movies show the position of individual cells within the population. These images and movies can be directly compared with the predictions of discrete models, whereas continuum models do not provide direct information about individuals within the population [29–31].

Discrete time, lattice-based random walk models of collective cell migration assays typically invoke the following discrete proliferation mechanism [23–26]. During any time step, of duration τ, each agent is given the chance to proliferate with probability P_p. To incorporate crowding effects, the target site for the placement of the daughter agent is chosen, at random, from the nearest neighbouring lattice

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

sites surrounding the mother agent. If the target site is vacant, a daughter agent will be placed onto the target site. However, if the target site is occupied, the event is aborted. Mean field analysis of this kind of proliferation mechanism shows that it gives rise to a logistic source term in the partial differential equation (PDE) description of the model [8, 25, 26]. The logistic model is the most commonly-invoked source term in continuum reaction-diffusion models of collective cell migration assays [5, 6, 12–19].

Although carrying-capacity limited proliferation has been traditionally modelled using the logistic equation, $\mathrm{d}C/\mathrm{d}t = \lambda C(1-C)$ [5, 6, 12–15, 19], where λ is the proliferation rate and the density has been scaled relative to the carrying capacity density, there is a growing awareness that the logistic model does not always match experimental data. For example, West and colleagues [32] compare data describing the growth of tumour spheroids in a wide range of animal models and show that the growth is best described by a generalised logistic model, $\mathrm{d}C/\mathrm{d}t = \lambda C^{3/4}(1-C^{1/4})$. Similarly, our previous analysis of a suite of scratch assays suggests that when a logistic-type reaction-diffusion equation is calibrated to match experimental data with a range of initial cell densities, there is no unique choice of λ for which the logistic model matches the entire data set for all initial conditions considered [10]. One way of interpreting this result is that cells do not proliferate logistically. While several previous theoretical studies have analysed generalised logistic growth models, such as $\mathrm{d}C/\mathrm{d}t = \lambda C^{\alpha}(1-C^{\beta})^{\gamma}$ for arbitrary constants α, β and γ [33], the question of which discrete exclusion process mechanisms correspond to these generalised logistic growth models is yet to be examined.

The aim of this work is to present, and analyse, a discrete model of two-dimensional collective cell migration assays. In all cases we consider the cell migration to be modelled in a standard way, namely as an unbiased, nearest neighbour exclusion process where potential migration events occur with probability P_m per time step of duration τ [8]. The focus of our work is on the details of the proliferation mechanism. Potential proliferation events occur with probability P_p per time step of duration τ. In the traditional model, the location of the daughter agent is chosen by randomly selecting a nearest neighbour lattice site. If the randomly selected target site is vacant, the proliferation event is successful, whereas if the randomly selected target site is occupied, the proliferation event is aborted [8, 25, 26]. Two extensions of the standard model are analyzed. (i) We first consider non-nearest neighbour proliferation mechanisms, whereby the crowdedness of any individual agent is influenced by a larger template on the lattice, and it is possible for the daughter agent to be placed on a non-nearest neighbour site. This first extension is consistent with *in vitro* [34] and *in vivo* [35] experimental observations of non-nearest neighbour proliferation mechanisms. (ii) We note that the traditional discrete proliferation mechanism corresponds to a linear relationship between the per capita growth rate and the density, $(1/C)\mathrm{d}C/\mathrm{d}t = \lambda(1-C)$ [8, 25, 26]. By adjusting the way that we measure the local density, $\hat{C} \in [0,1]$, we implement a suite of discrete models that correspond to $(1/\hat{C})\mathrm{d}\hat{C}/\mathrm{d}t = \lambda f(\hat{C})$, by introducing a *crowding function*, $f(C) \in [0,1]$ with $f(0) = 1$ and $f(1) = 0$.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The focus of this work is to design and implement a generalised cell proliferation mechanism into a lattice-based random walk model of cell migration and proliferation that incorporates these two extensions. We derive the continuum limit PDE description of the discrete model, and apply both the discrete and continuum models to mimic a suite of cell proliferation and scratch assays. Our results illustrate several interesting features about the relationship between the discrete model and the continuum limit description.

II. DISCRETE MATHEMATICAL MODELS

We adopt the convention that dimensional variables are primed, whereas non-dimensional variables are unprimed. A lattice-based random walk model will be used to describe the collective motion of a population of cells with an average cell diameter of Δ'. The lattice spacing is taken to be equal to the average cell diameter so that there are, at most, one agent per site. All simulations are non-dimensional in the sense that they are performed on a hexagonal lattice with unit lattice spacing, $\Delta = 1$. These non-dimensional simulations can be used to model any particular cell population by re-scaling with the dimensional cell diameter, Δ'.

Each lattice site, indexed (i, j) where $i, j \in \mathbb{Z}^+$, has position

$$(x, y) = \begin{cases} \big(i\Delta, j\Delta\sqrt{3}/2\big) & \text{if } j \text{ is even,} \\ \big((i + 1/2)\Delta, j\Delta\sqrt{3}/2\big) & \text{if } j \text{ is odd,} \end{cases}$$

such that $1 \leq i \leq I$ and $1 \leq j \leq J$. In any single realisation of the model, the occupancy of site $\mathbf{s}$ is denoted $C_{\mathbf{s}}$, with $C_{\mathbf{s}} = 1$ if the site is occupied, and $C_{\mathbf{s}} = 0$ if vacant. Since site $\mathbf{s}$ is associated with a unique index (i, j), we will use $C_{\mathbf{s}}$ and $C_{i,j}$ interchangeably.

Traditional discrete model: If there are $N(t)$ agents present at time t, then during the next time step of duration τ, $N(t)$ agents are selected independently at random, one at a time with replacement, and given the opportunity to move [8, 26]. The randomly selected agent attempts to move, with probability P_m, to one of the six nearest neighbour sites (Figure 2(a)), and the target site is chosen randomly. To mimic crowding effects, a motility event is aborted if an agent attempts to move to an occupied site.

Once $N(t)$ potential motility events are attempted, another $N(t)$ agents are selected independently, at random, one at a time with replacement, and given the opportunity to proliferate with probability P_p. The location of the daughter agent is chosen, at random, from one of the six nearest neighbour sites [8, 25, 26]. If the selected site is occupied, the potential proliferation event is aborted. In contrast, if the selected site is vacant, a new daughter agent is placed on that site. After the $N(t)$ potential

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

proliferation events have been attempted, $N(t+\tau)$ is updated [8, 25, 26]. One of the limitations of the traditional discrete model is that the continuum limit description leads to the traditional logistic source term [25, 26], however several observations indicate that logistic growth is not always appropriate [10, 32]. As such, we consider two extensions of the discrete proliferation mechanism to explore how to incorporate more general, universal, growth functions.

Extension 1: We first generalise the traditional discrete proliferation mechanism so that crowding effects are felt over a larger spatial template, and daughter agents can be placed on non-nearest neighbour lattice sites. The placement of daughter agents on non-nearest neighbour sites is consistent with previous *in vitro* [34] and *in vivo* [35] experimental observations. To achieve this goal, we consider a proliferative agent at site $\mathbf{s}$, and we use $\mathcal{N}_r\{\mathbf{s}\}$ to denote the set of neighbouring sites, where $r \geq 1$ is the number of concentric rings of lattice sites surrounding $\mathbf{s}$. For example, when $r = 1$, $\mathcal{N}_1\{\mathbf{s}\}$ denotes the set of six nearest-neighbouring sites, as demonstrated in Figure 2(a). In contrast, when $r = 2$, $\mathcal{N}_2\{\mathbf{s}\}$ also includes the set of the next nearest-neighbouring sites, as demonstrated in Figure 2(e). More generally, the number of sites in $\mathcal{N}_r\{\mathbf{s}\}$ is $Z_{\mathcal{N}_r} = 3r(r+1)$.

To implement this extension we first choose a value of $r \geq 1$. For any potential proliferation event, the target site for the placement of the daughter agent is chosen from $\mathcal{N}_r\{\mathbf{s}\}$. If a randomly chosen target site is vacant, a daughter agent is placed at that site. If a randomly chosen target site is occupied the potential proliferation event is aborted.

Extension 2: Instead of deciding to abort a potential proliferation event depending on the occupancy of a single randomly chosen site, we now consider a more general approach to incorporate crowding effects. We assume that a proliferative agent at site $\mathbf{s}$ senses the occupancy of all sites within $\mathcal{N}_r\{\mathbf{s}\}$, and detects a measure of the average occupancy of those sites, $\hat{C}_{\mathbf{s}} = (1/Z_{\mathcal{N}_r}) \sum_{\mathbf{s}' \in \mathcal{N}_r\{\mathbf{s}\}} C_{\mathbf{s}'}$. This means that $\hat{C}_{\mathbf{s}} \in [0, 1]$ is a measure of the crowdedness of the region surrounding $\mathbf{s}$. We anticipate that using $\hat{C}_{\mathbf{s}}$ to determine whether potential proliferation events are aborted is more realistic than the traditional model where the decision depends solely on the occupancy of a single, randomly chosen site.

To use $\hat{C}_{\mathbf{s}}$ to determine whether a potential proliferation event succeeds, we introduce a *crowding function*, $f(C) \in [0, 1]$ with $f(0) = 1$ and $f(1) = 0$. To incorporate crowding effects we sample a random number, $R \sim U(0, 1)$. If $R < f(\hat{C}_{\mathbf{s}})$, a daughter agent is placed at a randomly chosen vacant site in $\mathcal{N}_r\{\mathbf{s}\}$, whereas if $R > f(\hat{C}_{\mathbf{s}})$, the potential proliferation event is aborted. This extension can be applied to different sized templates by varying $r \geq 1$.

Generalised discrete model: We now implement an algorithm that can be used to simulate both extensions 1 and 2. In this generalised model, a proliferative agent can place a daughter agent at any vacant target site in $\mathcal{N}_r\{\mathbf{s}\}$, and crowding effects are modeled by $f(C)$. Therefore, during a potential proliferation event, a randomly selected agent at site $\mathbf{s}$ attempts to proliferate with probability P_p per

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

time step of duration τ. If the agent is to attempt to proliferate, crowding effects are incorporated by calculating $f(\hat{C}_{\mathbf{s}})$. If the potential proliferation event is to succeed, a daughter agent is placed at a randomly selected vacant site in $\mathcal{N}_r\{\mathbf{s}\}$.

Figure 2 shows four different schematic illustrations of the generalised proliferation mechanism on a hexagonal lattice. In each illustration, three examples are shown in which sites in $\mathcal{N}_r\{\mathbf{s}\}$ are either: (i) all vacant with $\hat{C}_{\mathbf{s}} = 0$ (Figure 2(a), (e), (i) and (m)); (ii) half occupied with $\hat{C}_{\mathbf{s}} = 0.5$ (Figure 2(b), (f), (j) and (n)); or, (iii) fully occupied with $\hat{C}_{\mathbf{s}} = 1$ (Figure 2(c), (g), (k) and (o)). The first row (Figure 2(a)-(d)) corresponds to the traditional model with $r = 1$ and $f(C) = 1 - C$. In this case, with $\hat{C}_{\mathbf{s}} = 0.5$, the probability that a potential proliferation event takes place is $P_p/2$, and the probability of a particular proliferation event producing a daughter agent at a particular site is $P_p/6$ (Figure 2(b)). The second row (Figure 2(e)-(h)) corresponds to $r = 2$ and $f(C) = 1 - C$. In this case, with $\hat{C}_{\mathbf{s}} = 0.5$, the probability that a potential proliferation event takes place is $P_p/2$, and the probability of a particular proliferation event producing a daughter agent at a particular site is $P_p/18$ (Figure 2(f)). Comparing the outcomes in the first and second row of Figure 2 illustrates the first generalisation of the discrete proliferation mechanism as we are simply applying the same proliferation mechanism, with the same crowding function, over a larger template of lattice sites.

The schematic illustrations in the third (Figure 2(i)-(l)) and fourth (Figure 2(m)-(p)) rows of Figure 2 show how the outcomes in the first and second rows can be generalised by choosing different $f(C)$. For example, with $f(C) = (1-C)^2$, agents are less likely to proliferate than when $f(C) = 1-C$. With $r = 1$, $\hat{C}_{\mathbf{s}} = 0.5$ and $f(C) = (1-C)^2$ (Figure 2(j)), the probability that a potential proliferation event takes place is $P_p/4$, and the probability of a particular proliferation event producing a daughter agent at a particular site is $P_p/12$, and this is very different to the traditional model with $r = 1$, $\hat{C} = 0.5$ and $f(C) = 1 - C$ (Figure 2(b)). Similarly, with $r = 2$, $\hat{C}_{\mathbf{s}} = 0.5$ and $f(C) = (1 - C)^2$ (Figure 2(n)), the probability that a potential proliferation event takes place is $P_p/4$, and the probability of a particular proliferation event producing a daughter agent at a particular site is $P_p/36$, which is very different to the case where $f(C) = 1 - C$ with $r = 2$ and $\hat{C}_{\mathbf{s}} = 0.5$ (Figure 2(f)).

III. CONTINUUM DESCRIPTION

While the individual-level details of the generalised discrete proliferation mechanism, highlighted in Figure 2, are very different to the traditional proliferation mechanism, it is not immediately clear how these differences manifest when we consider the collective behaviour of a population of motile and proliferative agents. To investigate this question we will first derive the mean field continuum limit description of the discrete models, and then compare the performance of the continuum limit descriptions with averaged data from repeated simulations of the discrete model.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Traditional model: We first explain how the continuum limit description of the traditional model can be derived before we consider the more general case. We average the occupancy of site $\mathbf{s}$ over many identically prepared realisations to obtain $\langle C_{\mathbf{s}}\rangle \in [0,1]$ [26] and then develop a discrete conservation statement describing the change in average occupancy of site $\mathbf{s}$ from time t to time $t+\tau$,

$$\delta\langle C_{\mathbf{s}}\rangle = \frac{P_m}{6}\left(1-\langle C_{\mathbf{s}}\rangle\right)\sum_{\mathbf{s}'\in\mathcal{N}_1\{\mathbf{s}\}}\langle C_{\mathbf{s}'}\rangle - \frac{P_m}{6}\langle C_{\mathbf{s}}\rangle\left(6-\sum_{\mathbf{s}'\in\mathcal{N}_1\{\mathbf{s}\}}\langle C_{\mathbf{s}'}\rangle\right) + \frac{P_p}{6}\left(1-\langle C_{\mathbf{s}}\rangle\right)\sum_{\mathbf{s}'\in\mathcal{N}_1\{\mathbf{s}\}}\langle C_{\mathbf{s}'}\rangle, \quad (1)$$

where $\sum_{\mathbf{s}'\in\mathcal{N}_1\{\mathbf{s}\}}\langle C_{\mathbf{s}'}\rangle$ is the sum of the average occupancy of sites in $\mathcal{N}_1\{\mathbf{s}\}$. The first and second terms on the right of Equation (1) represent the effects of migration into, and out of, site $\mathbf{s}$, respectively. The third term on the right of Equation (1) represents the effect of proliferation. To arrive at Equation (1), we make the usual mean field assumption that the occupancy of lattice sites are independent as we interpret the products of terms like $\langle C_{\mathbf{s}}\rangle$ and $\sum_{\mathbf{s}'\in\mathcal{N}_1\{\mathbf{s}\}}\langle C_{\mathbf{s}'}\rangle$ as a net transition probability [25, 26, 36, 37].

We expand each term in Equation (1) as a Taylor series about site $\mathbf{s}$, neglect terms of $\mathcal{O}(\Delta^3)$, and divide both sides of the resulting expression by τ. Identifying $\langle C_{\mathbf{s}}\rangle$ with a smooth function, $C(x,y,t)$, we consider the limit as $\Delta\to 0$ and $\tau\to 0$ jointly, with the ratio Δ^2/τ held constant. This process leads to the PDE [26]

$$\frac{\partial C}{\partial t} = D\nabla^2 C + \lambda C\left(1-C\right), \quad (2)$$

where the diffusivity is $D = (P_m/4)\lim_{\Delta\to 0,\tau\to 0}\left(\Delta^2/\tau\right)$, and the proliferation rate is $\lambda = \lim_{\tau\to 0}\left(P_p/\tau\right)$. To obtain a well-defined continuum limit we require $P_p = \mathcal{O}(\tau)$ [21, 26]. Equation (2) confirms that the traditional proliferation mechanism is associated with the usual logistic source term. Furthermore, in one dimension, Equation (2) simplifies to the Fisher-Kolmogorov model [38, 39].

Generalised model: The migration mechanism in the traditional and generalised models are equivalent, whereas the generalised proliferation mechanism involves both a variable size template and a different mechanism, based on $f(C)$, to incorporate crowding. The corresponding discrete conservation statement is

$$\delta\langle C_{\mathbf{s}}\rangle = \frac{P_m}{6}\left(1-\langle C_{\mathbf{s}}\rangle\right)\sum_{\mathbf{s}'\in\mathcal{N}_1\{\mathbf{s}\}}\langle C_{\mathbf{s}'}\rangle - \frac{P_m}{6}\langle C_{\mathbf{s}}\rangle\left(6-\sum_{\mathbf{s}'\in\mathcal{N}_1\{\mathbf{s}\}}\langle C_{\mathbf{s}'}\rangle\right) + \frac{P_p}{Z_{\mathcal{N}_r}}\left(1-\langle C_{\mathbf{s}}\rangle\right)\sum_{\mathbf{s}'\in\mathcal{N}_1\{\mathbf{s}\}}\langle C_{\mathbf{s}'}\rangle\frac{f\left(\langle\hat{C}_{\mathbf{s}'}\rangle\right)}{1-\langle\hat{C}_{\mathbf{s}'}\rangle}, \quad (3)$$

where $\langle\hat{C}_{\mathbf{s}}\rangle = (1/Z_{\mathcal{N}_r})\sum_{\mathbf{s}'\in\mathcal{N}_r\{\mathbf{s}\}}\langle C_{\mathbf{s}'}\rangle$. The first two terms on the right of Equation (3) are identical to the corresponding terms in Equation (1). The third term on the right of Equation (3) represents the change in occupancy of site $\mathbf{s}$ due to proliferation. The factor $f(\langle\hat{C}_{\mathbf{s}'}\rangle)/(1-\langle\hat{C}_{\mathbf{s}'}\rangle)$ is a measure of the crowding at site $\mathbf{s}'$, in terms of $f(C)$, relative to the probability that sites in the neighbourhood are vacant. Again,

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Equation (3) implicitly assumes that the occupancy of lattice sites are independent [25, 36, 37]. To obtain a continuum description, we expand each term in Equation (3) in a Taylor series about site $\mathbf{s}$ and truncate terms of $\mathcal{O}(\Delta^3)$. Dividing the resulting expression by τ, we identify $\langle C_{\mathbf{s}} \rangle$ with a smooth function, $C(x, y, t)$, and consider the limit as $\Delta \to 0$ and $\tau \to 0$ jointly, with the ratio Δ^2/τ held constant. The resulting PDE is

$$\frac{\partial C}{\partial t} = D\nabla^2 C + \lambda C f\left(C\right). \tag{4}$$

Details of the Taylor series expansions used to derive Equation (4) are given in Appendix A. Comparing Equations (2) and (4), the population-level impact of the change in the proliferation mechanism is to alter the per capita growth rate from the linearly decreasing function of density, $\lambda(1-C)$, to the more general $\lambda f(C)$. Therefore, the generalised discrete mechanism encodes more general crowding effects than the traditional model. For the remainder of this work we set $f(C) = C^{\alpha-1}(1-C^{\beta})^{\gamma}$, where α, β and γ are constants [33], but many other choices of $f(C)$ are possible.

IV. RESULTS AND DISCUSSION

Our main result, so far, is to describe how to incorporate a generalised proliferation mechanism into a discrete two-dimensional model of cell migration and cell proliferation with crowding effects, and to derive the mean field continuum limit description. However, at this stage, it is unclear how well the continuum model will predict averaged data from repeated stochastic simulations of the discrete model. To explore this issue we now apply the discrete and continuum models to mimic both a cell proliferation assay and a scratch assay (Figure 1). We systematically vary r and $f(C)$ to explore how these choices affect the performance of the continuum description.

A key parameter in all applications of the discrete model is P_p/P_m, which measures the frequency of proliferation events relative to the frequency of motility events for an isolated agent [26]. This ratio can be estimated from experimentally observable quantities including the doubling time $t_d' = (\tau' \log_e 2)/P_p$, the cell diffusivity D', and the average cell diameter, Δ'. For typical values: $t_d' = 24$h [26]; $D' = 1000$ μm^2/h [8]; and, $\Delta' = 24\mu$m [10], we have $P_p/P_m \approx 0.001$. Therefore, all simulations and analysis in the main document correspond to $\Delta = \tau = P_m = 1$ and $P_p = 0.001$. These non-dimensional simulations can be used to model a population of cells with an arbitrary dimensional cell diameter and an arbitrary doubling time by re-scaling Δ and τ with appropriate choices of Δ' and τ', respectively [26]. To ensure the conclusions drawn from these simulations are applicable to a wide range of cell lines, we repeat all simulations and analysis for a higher proliferation rate, $\Delta = \tau = P_m = 1$ and $P_p = 0.05$ (Supplementary Material). Therefore, the results in the main document correspond to a typical cell line with a standard doubling time of $t_d' = 24$h, while additional results correspond to a much smaller

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

doubling time of $t_d' = 14$h (Supplementary Material).

Cell proliferation assay. We consider a suite of simulations of a cell proliferation assay based on the geometry of the images in Figure 1(a)-(c). We use a lattice of size $I \times J$ to accommodate a typical population of cells ($\Delta' = 24\mu$m [10]). Since the images in Figure 1(a)-(c) show a fixed field of view that is much smaller than the spatial extent of the real experiment, and the cells in the experiment are distributed uniformly, we apply zero net flux boundary conditions along all boundaries [4]. Simulations are initiated by randomly populating each lattice site with a constant probability, so that there is, on average, no spatial gradients in agent density across the lattice. For each combination of r and $f(C)$, we consider two different initial densities: 5% and 20%. Snapshots from the model are given in Figure 3.

To mimic the way that cell proliferation assays are reported [7], we calculate the time evolution of the total number of agents on the lattice, which, when divided by the total number of lattice sites, gives the agent density per unit area [26]

$$\langle C \rangle = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} C_{i,j}.$$

We further average these results over many identically prepared simulations so that we report relatively smooth data where stochastic fluctuations in the agent density are negligible. To compare averaged simulation data with the solution of the continuum model, we note that the absence of spatial gradients in the cell proliferation assay means that, on average, $\nabla^2 C = 0$. Therefore, instead of dealing with a PDE for $C(x, y, t)$, Equation (4) simplifies to the ordinary differential equation (ODE) for $C(t)$ [7]

$$\frac{\mathrm{d}C}{\mathrm{d}T} = Cf(C), \tag{5}$$

where we have written $T = t\lambda$. This re-scaling of the time variable allows us to more easily compare results in the main paper for a standard proliferation rate ($P_p/P_m = 0.001$) with additional results for faster proliferation ($P_p/P_m = 0.05$) (Supplementary Material). Although it is possible to solve Equation (5) exactly for certain choices of $f(C)$, it is not always possible to find an explicit solution for other relevant choices of $f(C)$. Therefore we solve Equation (5) numerically using a backward Euler approximation with a constant time step, δt, and Picard linearisation with convergence tolerance, ϵ [40].

Results in Figure 4 show that when $f(C)$ is linear, the discrete and continuum density profiles are indistinguishable for both the lower (Figure 4(b)) and higher (Figure 4(c)) initial density simulations. Results in Figure 4(d)-(e) quantify the discrepancy between the solution of the continuum model and averaged discrete density data using $E = \langle C \rangle - C$, where $\langle C \rangle$ is the average density per unit area from the discrete simulations and C is the solution of Equation (5). In summary, the evolution of E

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(Figure 4(d)-(e)) shows that the error is extremely small, with no discernible trends for the different choices of r.

Results in Figures 5-7 show similar results to those in Figure 4, except now we consider a range of nonlinear $f(C)$. The averaged discrete data and the solution of the corresponding continuum model show that, broadly speaking, the continuum model provides a good prediction of the averaged discrete results. Furthermore, the same high quality of match between the solution of the continuum models and the averaged discrete data is observed for both the lower and higher initial densities for all choices of $f(C)$ considered. When we compare the evolution of the density data between Figures 4-7, we see that the choice of $f(C)$ has a profound effect on the evolution of the experiment. For example, in cases where $f(C)$ is larger than the traditional linear crowding function, such as in Figure 5, we see that the growth process is faster than in the traditional case, such as Figure 4.

Although the match between the average discrete data and the solution of the corresponding continuum model in Figures 5-7 is very good, there are some trends that are not obvious without making these comparisons explicit. For example, results in Figures 5-7(b)-(c) indicate that the performance of the continuum model is slightly poorer than in Figure 4 where $f(C)$ is linear. However, in all cases, the performance of the continuum model improves as r increases. For example, all results with $r = 4$ lead to an excellent match regardless of $f(C)$. Therefore, these results indicate that estimating r from experimental time lapse images will be important if we need to decide whether the continuum approximation is sufficient, or whether we need to use more computationally demanding repeated simulations of the discrete model to accurately predict the average behaviour of the experiment. Similar results and trends are observed when the proliferation rate is much higher (Supplementary Material).

Scratch assay. We also consider a suite of simulations of a scratch assay based on the geometry of the images in Figure 1(d)-(f). Since the images in Figure 1(d)-(f) show a fixed field of view that is much smaller than the spatial extent of the real experiment [10], we apply zero net flux boundary conditions along all boundaries of the lattice. We model the scratch assays on a lattice of size $I \times J$ that is chosen to accommodate a typical population of cells ($\Delta' = 24\mu$m [10]). To model the initial condition, we randomly populate all lattice sites with an equal probability of 30%. To simulate the scratch, we remove all agents from a vertical region, with a width of 23 cell diameters (Figure 1(d)). Snapshots from the discrete model, for a range of $f(C)$ is shown in Figure 8.

Since the initial condition is uniform in the vertical direction, we average the agent population density along each vertical column of lattice sites to obtain

$$\langle C_i \rangle = \frac{1}{J} \sum_{j=1}^{J} C_{i,j},$$

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

which is then further averaged over many identically prepared discrete simulations to reduce fluctuations in the averaged density profiles. This procedure allows us to plot the time evolution of the average agent density as a function of the horizontal coordinate. This is appropriate for our initial condition where the initial density is independent of the vertical coordinate [4, 10]. Typical results are shown in Figure 9, with the standard linear $f(C)$, and four different choices of r. In general, we see the effects of combined agent motility and agent proliferation since the agent density profiles moves into the initially-vacant region. The effects of proliferation can also be observed as the density profile increases with time, both within the initially-vacant region and at other locations behind the scratch.

To explore how well the continuum description matches this vertically-averaged discrete density data, we note that since the initial condition is independent of the vertical location (Figure 1(d)), we can average Equation (4) in the vertical direction [4, 10] to give,

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} + \lambda C f(C). \tag{6}$$

To solve Equation (6), we apply no flux boundary condition at both boundaries, and the initial condition is given by $C(x, 0) \equiv 0$ within the initially-vacant region, and $C(x, 0) \equiv 0.30$ outside of this region. We solve Equation (6) numerically by approximating the spatial derivative term with a central finite difference approximation with uniform spacing δx, and the temporal derivative term is approximated using a backward Euler method with a uniform time step, δt [40]. The resulting system of nonlinear algebraic equations is solved using Picard iteration with convergence tolerance, ϵ.

The numerical solution of Equation (6) is superimposed on the averaged discrete density profiles in Figure 9 for $f(C) = 1 - C$. As with the cell proliferation assay results in Figure 4, the quality of the continuum-discrete match is excellent for all r considered. Data in Figures 10-12 show similar results to those in Figure 9, except now we consider a range of different nonlinear $f(C)$. As with the results for the cell proliferation assay, we find that the averaged discrete data and the solution of the corresponding PDE confirms that the continuum model provides a good prediction of the averaged discrete results. Comparing the evolution of $C(x, T)$ between Figures 9-12, we see that the choice of $f(C)$ has a profound effect on the evolution of the scratch assay. For example, choices of $f(C)$ that are greater than the standard linear function lead to a faster healing response (Figure 10), whereas choices of $f(C)$ that are less than the standard linear function lead to slower healing (Figure 11). Again, as for the cell proliferation assays, overall we observe a good quality of match between the solution of the continuum model and the averaged agent density profiles across all choices of $f(C)$ and r for the scratch assays. However, for nonlinear $f(C)$, we observe some small discrepancies, and these discrepancies are most pronounced when $r = 1$. In contrast, the quality of the continuum-discrete match is excellent for larger r across all choices of $f(C)$ that we consider.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

V. CONCLUSIONS

In this work we present a new simulation tool to study two-dimensional collective cell spreading. Based on a two-dimensional exclusion process, we apply the discrete model to describe both a cell proliferation assay and a scratch assay, which are standard experiments used to study various applications including wound healing and malignant spreading [1, 2].

The main motivation of our work is the observation that most continuum [5, 6, 12–17] and discrete [8, 25, 26] models of collective cell spreading are associated with a logistic growth source term to model cell proliferation. Despite this, there is mounting *in vivo* [32] and *in vitro* [10] evidence suggesting that a logistic growth term does not always best describe the way that a cell population grows. Motivated by these observations, we extend the standard exclusion process based discrete model of cell proliferation. In the traditional model, each agent is given the opportunity to proliferate, and to assess the role of crowding effects, a potential target site for the deposition of a daughter agent is chosen, at random, from one of the nearest neighbouring sites. If the randomly chosen target site is occupied, the potential proliferation event is aborted. Alternatively, if the randomly chosen target site is vacant, the potential proliferation event succeeds. While this mechanism has been widely invoked, it seems biologically unreasonable to assume that a proliferative cell randomly inspects a single region of space and decides whether to proliferate, or not, depending on the occupancy of a single point in space.

Our new, generalised discrete model encompasses two extensions of the traditional model. The first extension involves choosing a variable proliferation template so that the target site is chosen from a set of sites contained within $r \geq 1$ concentric rings about the mother agent. The second extension involves a measure of the crowdedness of the mother agent, $\hat{C}_{\mathbf{s}} \in [0, 1]$. A *crowding function*, $f(C) \in [0, 1]$ with $f(0) = 1$ and $f(1) = 0$, which incorporates information from a group of neighbouring sites is used to quantify the influence of crowding. Analysing the mean field continuum limit of the generalised model shows that the usual logistic source term, $\lambda C(1 - C)$, is generalised to $\lambda C f(C)$. There are several interesting consequences of this mean field PDE, namely: (i) the traditional logistic source term corresponds to linear $f(C)$; (ii) the size of the template, r, does not appear in the continuum limit description; and, (iii) without making explicit comparisons, it is unclear how different choices of $f(C)$ and r affect the accuracy of the continuum limit description.

To provide insight into how different choices of $f(C)$ and r affect the accuracy of the continuum limit description, we generate averaged discrete data from the generalised random walk model applied to both a cell proliferation assay and a scratch assay for typical cell line with $P_p/P_m = 0.001$ [8]. Averaged simulation data are generated for a range of choices of $f(C)$ and r and we find that, overall, the continuum limit approximation provides a good prediction of the average behaviour of the stochastic simulations. While there is a modest discrepancy in the continuum-discrete match for nonlinear $f(C)$,

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

we find that the quality of the match improves as r increases. Therefore, to make a distinction between the need for using repeated stochastic simulations or simply working with the continuum description, we suggest that experimental time lapse images ought to be used to provide an estimate of r. To ensure that our conclusions are broadly applicable across a wide range of cell lines, all continuum-discrete comparisons are repeated for a cell line with a particularly fast proliferation rate $P_p/P_m = 0.05$ (Supplementary Material), and we find that the same general conclusions apply.

Since the focus of this work is to investigate the role of the proliferation mechanism, all our simulations and analysis invoke the most straightforward unbiased nearest neighbour exclusion motility mechanism. This mechanism is thought to be a good approximation of the motility of mesenchymal cell lines that are largely unaffected by cell-to-cell adhesion [8, 11]. However, if dealing with an epithelial cell line, it would be reasonable to invoke a different motility mechanism that involves cell-to-cell adhesion [36, 37, 41]. Under these conditions it would be interesting to extend the present analysis to investigate the performance of the continuum limit description of the generalised proliferation mechanism with these kinds of adhesive motility mechanisms.

Acknowledgments: This work is supported by the Australian Research Council (DP140100249, FT130100148). We thank Parvathi Haridas and Esha Shah for providing the images in Figure 1.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

VI. APPENDIX A

In this appendix we demonstrate how to derive Equation (3) with $r = 1$ and and arbitrary $f(C)$. Extending the derivation to deal with $r > 1$ is straightforward. To begin, we note that site $\mathbf{s}$, with position (x, y), has six nearest neighbouring sites: site $\mathbf{s}'_1$ with position $(x - \Delta, y)$; site $\mathbf{s}'_2$ with position $(x + \Delta, y)$; site $\mathbf{s}'_3$ with position $(x - \Delta/2, y + \Delta\sqrt{3}/2)$; site $\mathbf{s}'_4$ with position $(x + \Delta/2, y + \Delta\sqrt{3}/2)$; site $\mathbf{s}'_5$ with position $(x - \Delta/2, y - \Delta\sqrt{3}/2)$; and, site $\mathbf{s}'_6$ with position $(x + \Delta/2, y - \Delta\sqrt{3}/2)$. It is useful to first write down expressions for the average density of these six nearest neighbouring sites expanded in a Taylor series about (x, y),

$$\langle C_{\mathbf{s}'_1}\rangle = \langle C_{\mathbf{s}}\rangle - \frac{\partial\langle C_{\mathbf{s}}\rangle}{\partial x}\Delta + \frac{\partial^2\langle C_{\mathbf{s}}\rangle}{\partial x^2}\frac{\Delta^2}{2} + \mathcal{O}(\Delta^3), \tag{7}$$

$$\langle C_{\mathbf{s}'_2}\rangle = \langle C_{\mathbf{s}}\rangle + \frac{\partial\langle C_{\mathbf{s}}\rangle}{\partial x}\Delta + \frac{\partial^2\langle C_{\mathbf{s}}\rangle}{\partial x^2}\frac{\Delta^2}{2} + \mathcal{O}(\Delta^3), \tag{8}$$

$$\langle C_{\mathbf{s}'_3}\rangle = \langle C_{\mathbf{s}}\rangle - \frac{\partial\langle C_{\mathbf{s}}\rangle}{\partial x}\frac{\Delta}{2} + \frac{\partial\langle C_{\mathbf{s}}\rangle}{\partial y}\frac{\sqrt{3}\Delta}{2} + \left[\frac{1}{4}\frac{\partial^2\langle C_{\mathbf{s}}\rangle}{\partial x^2} + \frac{3}{4}\frac{\partial^2\langle C_{\mathbf{s}}\rangle}{\partial y^2} - \frac{\sqrt{3}}{2}\frac{\partial^2\langle C_{\mathbf{s}}\rangle}{\partial x\partial y}\right]\frac{\Delta^2}{2} + \mathcal{O}(\Delta^3), \tag{9}$$

$$\langle C_{\mathbf{s}'_4}\rangle = \langle C_{\mathbf{s}}\rangle + \frac{\partial\langle C_{\mathbf{s}}\rangle}{\partial x}\frac{\Delta}{2} + \frac{\partial\langle C_{\mathbf{s}}\rangle}{\partial y}\frac{\sqrt{3}\Delta}{2} + \left[\frac{1}{4}\frac{\partial^2\langle C_{\mathbf{s}}\rangle}{\partial x^2} + \frac{3}{4}\frac{\partial^2\langle C_{\mathbf{s}}\rangle}{\partial y^2} + \frac{\sqrt{3}}{2}\frac{\partial^2\langle C_{\mathbf{s}}\rangle}{\partial x\partial y}\right]\frac{\Delta^2}{2} + \mathcal{O}(\Delta^3), \tag{10}$$

$$\langle C_{\mathbf{s}'_5}\rangle = \langle C_{\mathbf{s}}\rangle - \frac{\partial\langle C_{\mathbf{s}}\rangle}{\partial x}\frac{\Delta}{2} - \frac{\partial\langle C_{\mathbf{s}}\rangle}{\partial y}\frac{\sqrt{3}\Delta}{2} + \left[\frac{1}{4}\frac{\partial^2\langle C_{\mathbf{s}}\rangle}{\partial x^2} + \frac{3}{4}\frac{\partial^2\langle C_{\mathbf{s}}\rangle}{\partial y^2} + \frac{\sqrt{3}}{2}\frac{\partial^2\langle C_{\mathbf{s}}\rangle}{\partial x\partial y}\right]\frac{\Delta^2}{2} + \mathcal{O}(\Delta^3), \tag{11}$$

$$\langle C_{\mathbf{s}'_6}\rangle = \langle C_{\mathbf{s}}\rangle + \frac{\partial\langle C_{\mathbf{s}}\rangle}{\partial x}\frac{\Delta}{2} - \frac{\partial\langle C_{\mathbf{s}}\rangle}{\partial y}\frac{\sqrt{3}\Delta}{2} + \left[\frac{1}{4}\frac{\partial^2\langle C_{\mathbf{s}}\rangle}{\partial x^2} + \frac{3}{4}\frac{\partial^2\langle C_{\mathbf{s}}\rangle}{\partial y^2} - \frac{\sqrt{3}}{2}\frac{\partial^2\langle C_{\mathbf{s}}\rangle}{\partial x\partial y}\right]\frac{\Delta^2}{2} + \mathcal{O}(\Delta^3), \tag{12}$$

where we have truncated terms of $\mathcal{O}(\Delta^3)$. From this point on we drop the angle bracket notation. There are two terms on the right of Equation (3) that are associated with agent motility, and one term on the right of Equation (3) that is associated with agent proliferation. We will deal with these two different types of terms separately.

For agent migration, we need to deal with terms like $\sum_{\mathbf{s}'\in\mathcal{N}_1\{\mathbf{s}\}} C_{\mathbf{s}'}$, which can be obtained by summing Equations (7)-(12) to give

$$\sum_{\mathbf{s}'\in\mathcal{N}_1\{\mathbf{s}\}} C_{\mathbf{s}'} = 6C_{\mathbf{s}} + \left(\frac{\partial^2 C_{\mathbf{s}}}{\partial x^2} + \frac{\partial^2 C_{\mathbf{s}}}{\partial y^2}\right)\frac{3\Delta^2}{2} + \mathcal{O}(\Delta^3).$$

For agent proliferation, we need to deal with terms like $f(\hat{C}_{\mathbf{s}'_1}), f(\hat{C}_{\mathbf{s}'_2}), f(\hat{C}_{\mathbf{s}'_3}), \ldots, f(\hat{C}_{\mathbf{s}'_6})$. In this appendix we will show how to deal with $f(\hat{C}_{\mathbf{s}'_1})$ and we note that the other terms are a straightforward

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

extension of these results. First we expand the local density in a Taylor series about site $\mathbf{s}$

$$\begin{aligned}\hat{C}_{\mathbf{s}_1'} &= \frac{1}{6}\sum_{\mathbf{s}''\in\mathcal{N}_1\{\mathbf{s}'\}} C_{\mathbf{s}''} \\ &= C_{\mathbf{s}_1'} + \left(\frac{\partial^2 C_{\mathbf{s}_1'}}{\partial x^2} + \frac{\partial^2 C_{\mathbf{s}_1'}}{\partial y^2}\right)\frac{\Delta^2}{4} + \mathcal{O}(\Delta^3), \\ &= C_{\mathbf{s}} - \frac{\partial C_{\mathbf{s}}}{\partial x}\Delta + \left(3\frac{\partial^2 C_{\mathbf{s}}}{\partial x^2} + \frac{\partial^2 C_{\mathbf{s}}}{\partial y^2}\right)\frac{\Delta^2}{4} + \mathcal{O}(\Delta^3). \end{aligned} \tag{13}$$

We re-write Equation (13) as $\hat{C}_{\mathbf{s}_1'} = C_{\mathbf{s}} + \bar{C}$, where $\bar{C} = \mathcal{O}(\Delta)$. Therefore, the crowding function at site $\mathbf{s}_1'$ can be expanded in the following way,

$$\begin{aligned} f(\hat{C}_{\mathbf{s}_1'}) &= f(C_{\mathbf{s}} + \bar{C}), \\ &= f(C_{\mathbf{s}}) + \frac{\mathrm{d}f(C_{\mathbf{s}})}{\mathrm{d}C}\bar{C} + \frac{\mathrm{d}^2 f(C_{\mathbf{s}})}{\mathrm{d}C^2}\frac{\bar{C}^2}{2} + \mathcal{O}(\Delta^3). \end{aligned} \tag{14}$$

Using Equations (13)-(14) we obtain a series expansion for $f(\hat{C}_{\mathbf{s}_1'})$. Repeating this procedure for $f(\hat{C}_{\mathbf{s}_2'}), f(\hat{C}_{\mathbf{s}_3'}), \ldots, f(\hat{C}_{\mathbf{s}_6'})$, we expand all terms in Equation (3) that are associated with proliferation in a similar way. With all terms in Equation (3) expanded about (x, y), we obtain

$$\delta C_{\mathbf{s}} = P_m\left(\frac{\partial^2 C_{\mathbf{s}}}{\partial x^2} + \frac{\partial^2 C_{\mathbf{s}}}{\partial y^2}\right)\frac{\Delta^2}{4} + P_p C_{\mathbf{s}} f(C_{\mathbf{s}}). \tag{15}$$

Following the same procedure outlined in the main manuscript, Equation (15) leads to Equation (4).

---

[1] CC Liang, AY Park, JL Guan. Nat Protocols. **2** 329 (2007).

[2] R Riahi, Y Yang, DD Zhang, PK Wong. J Lab Autom. **17** 59 (2012).

[3] Y Kam, C Guess, L Estrada, B Weidow, V Quaranta. BMC Cancer. **8** 198 (2008).

[4] ST Johnston, ST Shah, LK Chopin, DLS McElwain, MJ Simpson. BMC Syst Biol. **9** 38 (2015).

[5] PK Maini, DLS McElwain, DI Leavesley. Tissue Eng. **10** 475 (2004).

[6] PK Maini, DLS McElwain, D Leavesley. Appl Math Lett. **17** 757 (2004).

[7] A Tremel, AQ Cai, N Tirtaatmadja, BD Hughes, G Stevens, KA Landman. Chem Eng Sci. **64** 247 (2009).

[8] MJ Simpson, BJ Binder, P Haridas, BK Wood, KK Treloar, DLS McElwain, RE Baker. Bull Math Biol. **75** 871 (2013).

[9] ST Johnston, MJ Simpson, DLS McElwain. J R Soc Interface. **11** 20140325 (2014).

[10] W Jin, ET Shah, CJ Penington, SW McCue, MJ Simpson. J Theor Biol. **390** 136 (2016).

[11] MJ Simpson, KK Treloar, BJ Binder, P Haridas, KJ Manton, DI Leavesley, DLS McElwain, RE Baker.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

J R Soc Interface. **10** 20130007 (2013).

[12] AQ Cai, KA Landman, BD Hughes. J Theor Biol. **245** 576 (2007).

[13] BG Sengers, CP Please, ROC Oreffo. J R Soc Interface. **4** 1107 (2007).

[14] U Savla, LE Olson, CM Waters. J Appl Physiol. **96** 566 (2004).

[15] H Sheardown, YL Cheng. Chem Eng Sci. **51** 4517 (1996).

[16] JF Hammond, DM Bortz. Appl Math Comput. **218** 2494 (2011).

[17] CW Curtis, Bortz DM. Phys Rev E. **86** 066108 (2012).

[18] DA Hormuth II, JA Weis, SL Barnes, MI Miga, EC Rericha, V Quaranta, TE Yankeelov. Phys. Biol. **12** 046006 (2015).

[19] M Shakeel, PC Matthews, RS Graham, SL Waters. Math Med Biol. **20** 21 (2013).

[20] BD Hughes. *Random Walks and Random Environments*. Volume 1 (Oxford University Press, 1995).

[21] EA Codling, MJ Plank, S Benhamou. J R Soc Interface. **5** 813 (2008).

[22] TM Liggett. *Interacting Particle Systems* (Springer, 2005).

[23] M Bramson, P Calderoni, A De Masi, P Ferrari, J Lebowitz, RH Schonmann. J Stat Phys. **45** 905 (1986).

[24] AR Kerstein. J Stat Phys. **45** 921 (1986).

[25] T Callaghan, E Khain, LM Sander, RM Ziff. J Stat Phys. **122** 909 (2006).

[26] MJ Simpson, KA Landman, BD Hughes. Physica A. **389** 3779 (2010).

[27] E Khain, M Katakowski, N Charteris, F Jiang, M Chopp. Phys Rev E. **86** 011904 (2012).

[28] N Charteris, E Khain. New J Phys. **16** 025002 (2014).

[29] ARA Anderson, MAJ Chaplain. Bull Math Biol. **60** 857 (1998).

[30] P Gerlee, ARA Anderson. Phys Biol. **12** 056001 (2015).

[31] DG Harvey, AG Fletcher, JM Osborne, J Pitt-Francis. Comput Phys Commn. **192** 130 (2015).

[32] GB West, JH Brown, BJ Enquist. Nature. **413** 628 (2001).

[33] A Tsoularis, J Wallace. Math Biosci. **179** 21 (2002).

[34] KA Landman, AQ Cai, BD Hughes. Bull Math Biol. **69** 2119 (2007).

[35] NR Druckenbrod, ML Epstein. Dev Biol. **287** 125 (2005).

[36] C Deroulers, M Aubert, M Badoual, B Grammaticos. Phys Rev E. **79** 031917 (2009).

[37] K Anguige, C Schmeiser. J Math Biol. **58** 395 (2009).

[38] RA Fisher. Ann Eugen. **7** 353 (1937).

[39] A Kolmogorov, I Petrovsky, N Piscounov. Mosc Univ Bull Math. **1** 1 (1937).

[40] KW Morton, DF Mayers. *Numerical Solution of Partial Differential Equations*. (Cambridge University Press, 2005).

[41] ST Johnston, MJ Simpson, RE Baker. Phys Rev E. **85** 051922 (2012).

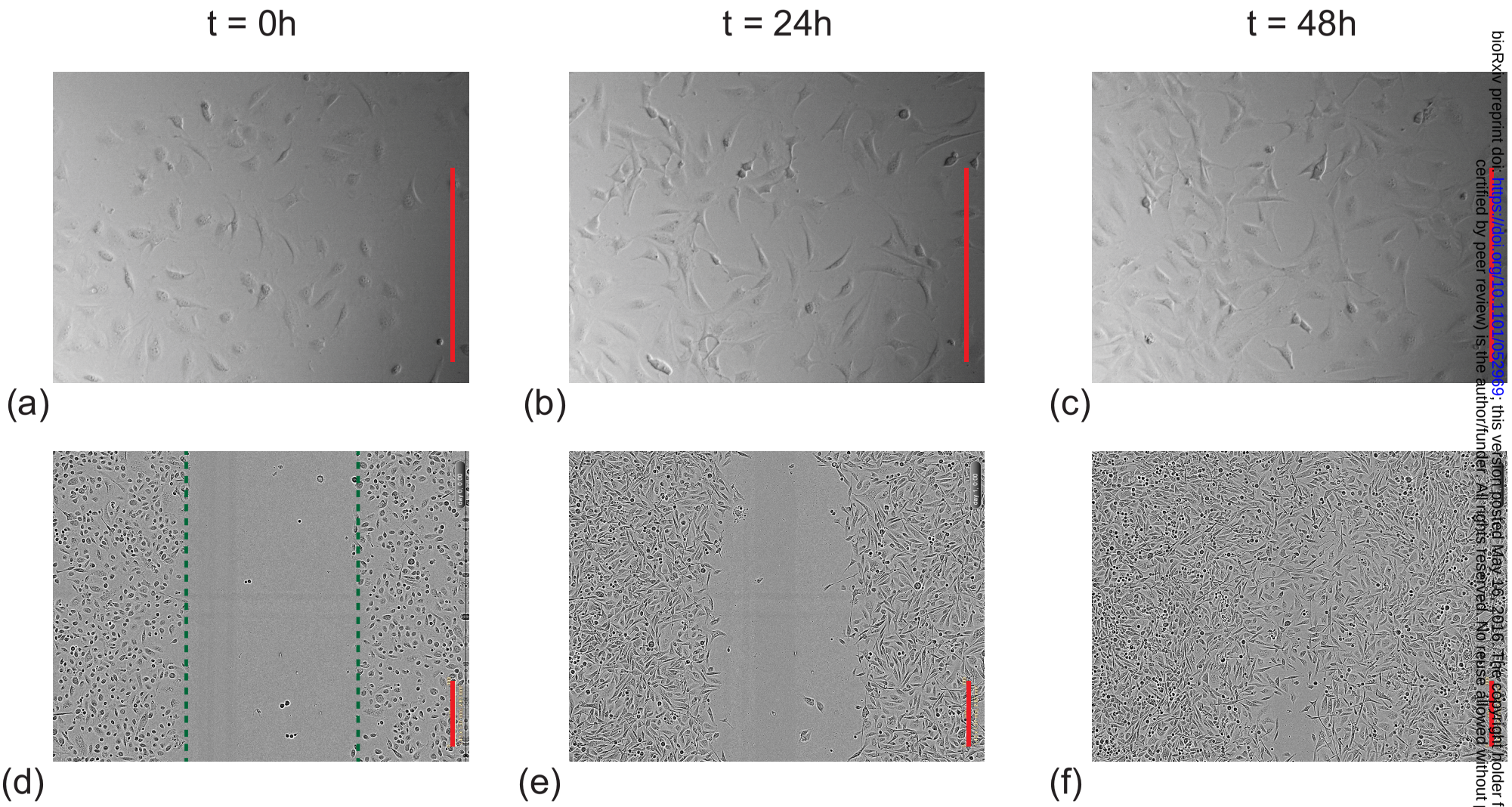

Figure 1. Experimental motivation. (a)-(c) Images from a *cell proliferation assay*, shown at $t = 0, 24$ and 48h, respectively [11]. The cell proliferation assay is initiated by uniformly distributing 25,000 3T3 fibroblast cells into the wells of a 24-well tissue culture plate [11]. The dimension of the field of view is $640\mu\text{m}\times480\mu\text{m}$, and the spatial extent of the growing population extends well beyond the field of view [11]. (d)-(f) Images from a *scratch assay*, shown at $t = 0, 24$ and 48h, respectively [10]. Experiments are initiated by uniformly distributing 16,000 PC3 cells into the well of a 96-well tissue culture plate [10]. A scratch (dashed green in (d)) is made at $t = 0$h [10], and the subsequent healing of the wound is observed with time. The dimension of the field of view is $1900\mu\text{m}\times1400\mu\text{m}$, and the spatial extent of the population extends well beyond the field of view [10]. The scale bar in each image corresponds to 300μm.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

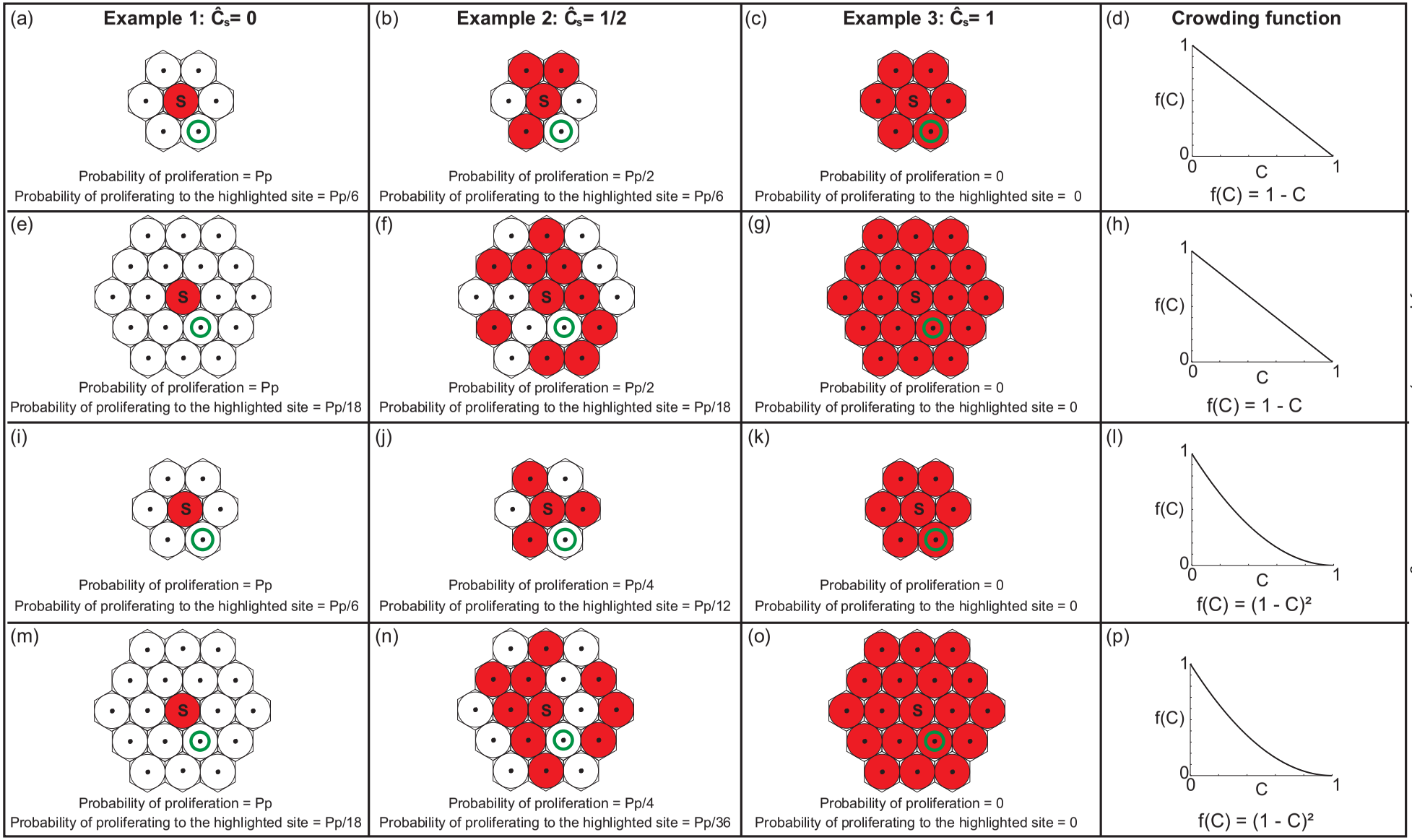

Figure 2. Schematic representation of the proliferation mechanisms considered in this work. Three examples are illustrated, in the first three columns, for $\hat{C} = 0, 0.5$ and 1, respectively. In each lattice fragment, the central site $\mathbf{s}$ is occupied (red), while some of the neighbouring sites are occupied (red) and others are vacant (white). In all cases we always consider the outcomes for a potential proliferation event of the agent at the central lattice site. The properties associated with the traditional proliferation mechanism are illustrated in (a)-(d). The properties associated with the first generalisation, where we consider the same crowding mechanism as the traditional model, but over a larger template of neighbouring lattice sites, is shown in (e)-(h). The properties associated with the second generalisation, where we consider the same nearest neighbour template as the traditional model, but we consider a different method of aborting potential proliferation events using $f(C)$, is shown in (i)-(l). The final row, (m)-(p), illustrates how the two generalisations can be combined. To highlight differences between the generalisations we report: $\hat{C}_{\mathbf{s}}$; the probability of a successful proliferation event taking place; and, the probability of a proliferation event successfully depositing a daughter at a particular site, highlighted with a green annulus, in (a)-(c), (e)-(g), (i)-(k) and (m)-(o).

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

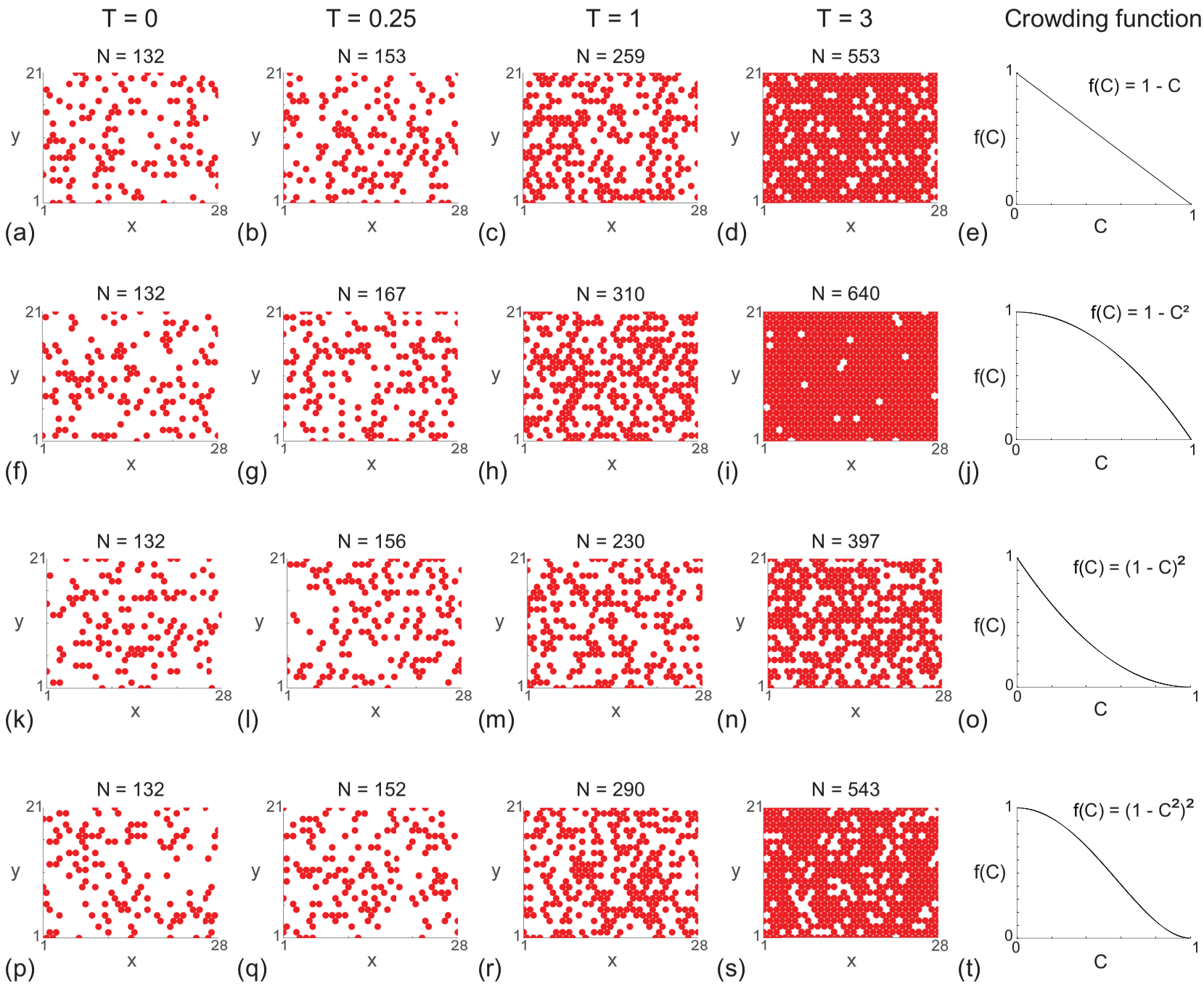

Figure 3. Snapshots of simulations for a suite of cell proliferation assays. In each row the distributions of agents at time $\lambda t = T = 0, 0.25, 1, 3$ are shown along with the corresponding $f(C)$. Each simulation is initiated by randomly populating a lattice of size $I = 28$ and $J = 24$, so that each site is occupied with probability 0.20. All simulations correspond to $\Delta = \tau = P_m = 1$, $P_p = 0.001$ and $r = 4$.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

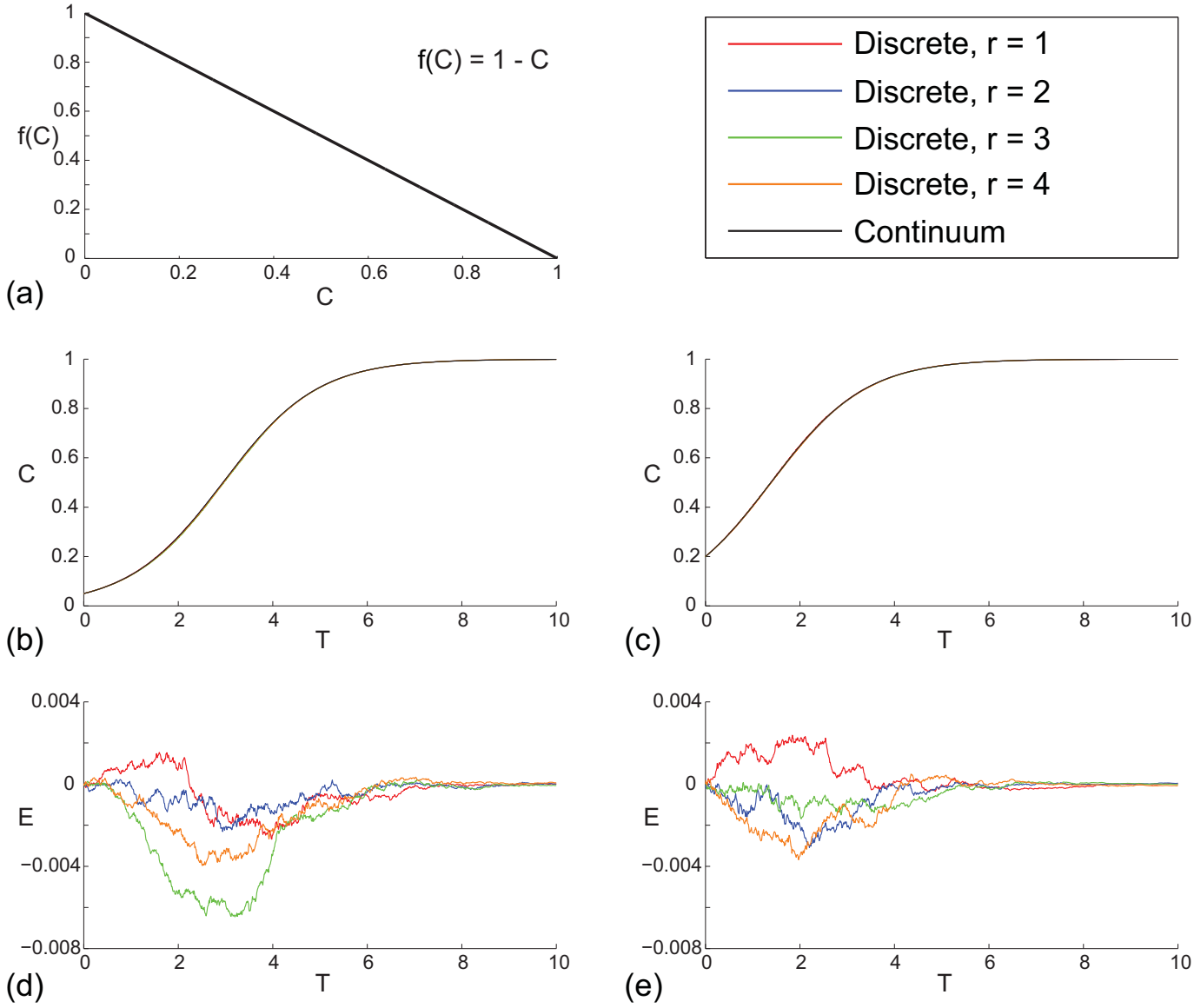

Figure 4. Comparison of averaged simulation data and the solution of the corresponding continuum model for a cell proliferation assay with $f(C) = 1 - C$, as shown in (a). Results in (b)-(c) compare averaged simulation data and the solution of the corresponding continuum model for a range of $1 \leq r \leq 4$. Two different initial conditions are considered so that in each realisation we have either 5% (b) or 20% (c) of lattice sites initially occupied. All simulation results are performed on a lattice with $I = 28$ and $J = 24$, and results are averaged across 300 identically prepared realisations of the discrete model. Profiles in (d)-(e) show the discrepancy between the solution of the continuum model and the average simulation data for the two different initial conditions shown in (b)-(c), respectively. All simulation results correspond to $\Delta = \tau = P_m = 1$ and $P_p = 0.001$, and the numerical solution of the continuum model is obtained with $\delta t = 1 \times 10^{-3}$ and $\epsilon = 1 \times 10^{-5}$.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

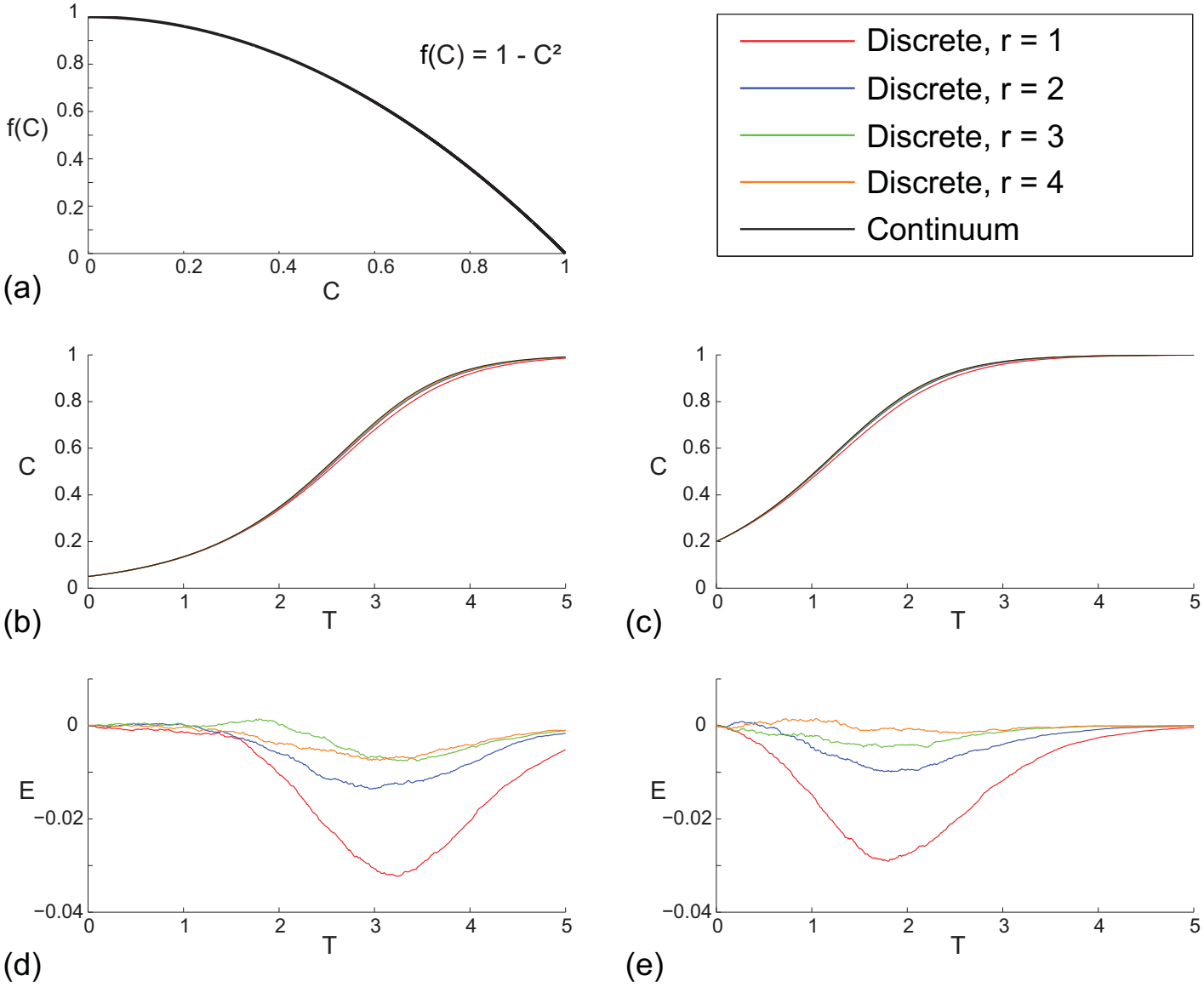

Figure 5. Comparison of averaged simulation data and the solution of the corresponding continuum model for a cell proliferation assay with $f(C) = 1 - C^2$, as shown in (a). Results in (b)-(c) compare averaged simulation data and the solution of the corresponding continuum model for a range of $1 \leq r \leq 4$. Two different initial conditions are considered so that in each realisation we have either 5% (b) or 20% (c) of lattice sites initially occupied. All simulation results are performed on a lattice with $I = 28$ and $J = 24$, and results are averaged across 300 identically prepared realisations of the discrete model. Profiles in (d)-(e) show the discrepancy between the solution of the continuum model and the average simulation data for the two different initial conditions shown in (b)-(c), respectively. All simulation results correspond to $\Delta = \tau = P_m = 1$ and $P_p = 0.001$, and the numerical solution of the continuum model is obtained with $\delta t = 1 \times 10^{-3}$ and $\epsilon = 1 \times 10^{-5}$.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

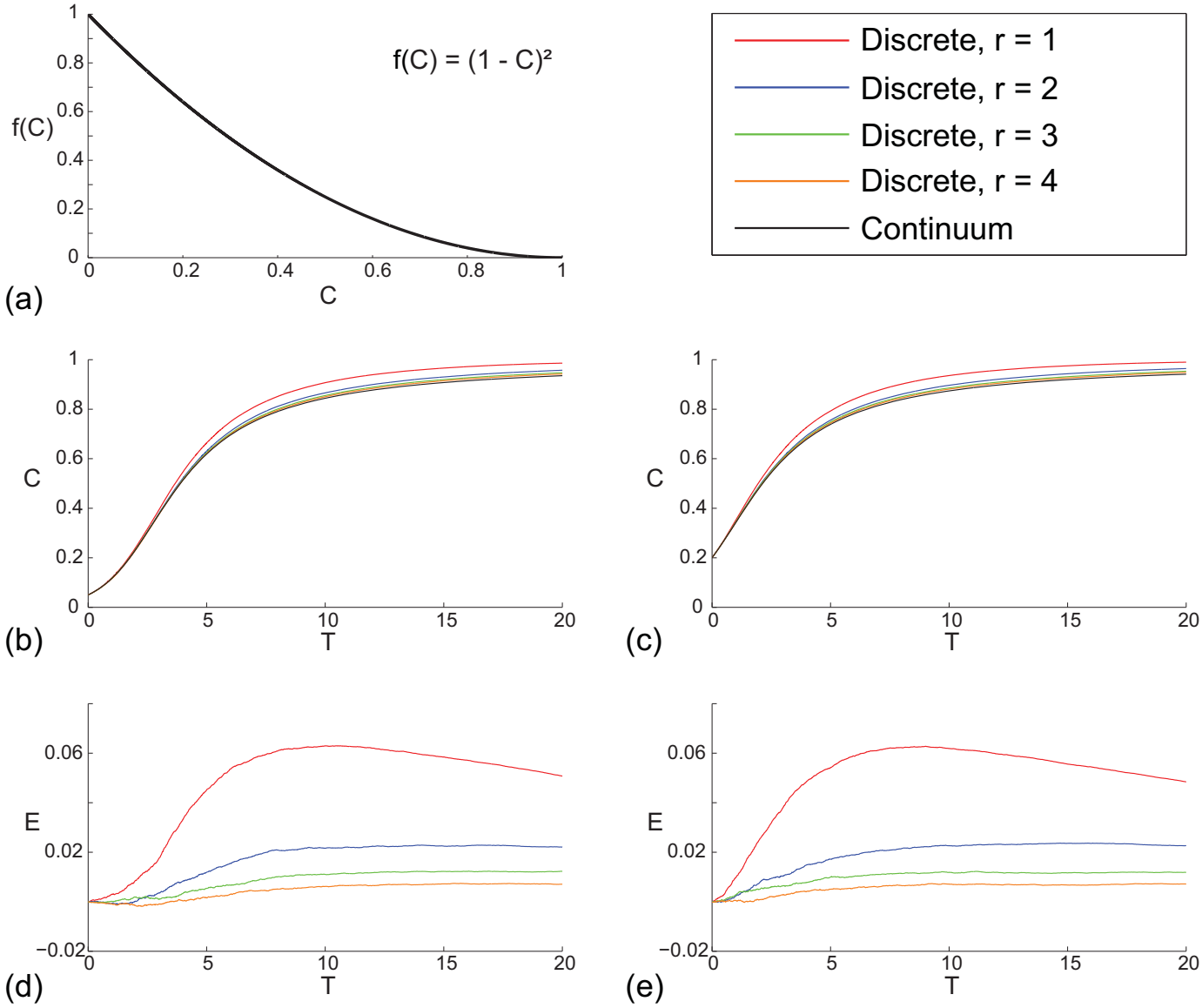

Figure 6. Comparison of averaged simulation data and the solution of the corresponding continuum model for a cell proliferation assay with $f(C) = (1 - C)^2$, as shown in (a). Results in (b)-(c) compare averaged simulation data and the solution of the corresponding continuum model for a range of $1 \leq r \leq 4$. Two different initial conditions are considered so that in each realisation we have either 5% (b) or 20% (c) of lattice sites initially occupied. All simulation results are performed on a lattice with $I = 28$ and $J = 24$, and results are averaged across 300 identically prepared realisations of the discrete model. Profiles in (d)-(e) show the discrepancy between the solution of the continuum model and the average simulation data for the two different initial conditions shown in (b)-(c), respectively. All simulation results correspond to $\Delta = \tau = P_m = 1$ and $P_p = 0.001$, and the numerical solution of the continuum model is obtained with $\delta t = 1 \times 10^{-3}$ and $\epsilon = 1 \times 10^{-5}$.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

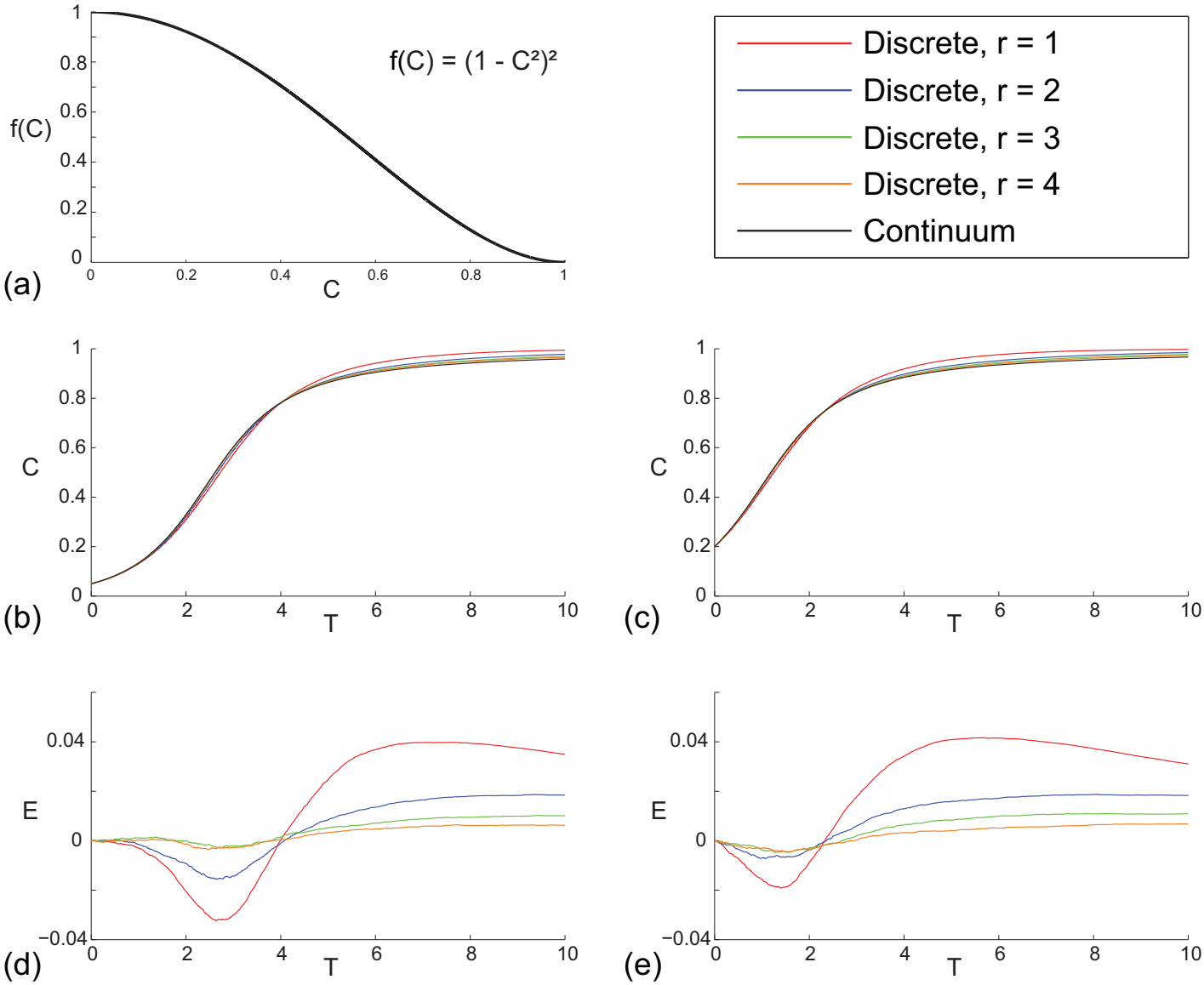

Figure 7. Comparison of averaged simulation data and the solution of the corresponding continuum model for a cell proliferation assay with $f(C) = (1 - C^2)^2$, as shown in (a). Results in (b)-(c) compare averaged simulation data and the solution of the corresponding continuum model for a range of $1 \leq r \leq 4$. Two different initial conditions are considered so that in each realisation we have either 5% (b) or 20% (c) of lattice sites initially occupied. All simulation results are performed on a lattice with $I = 28$ and $J = 24$, and results are averaged across 300 identically prepared realisations of the discrete model. Profiles in (d)-(e) show the discrepancy between the solution of the continuum model and the average simulation data for the two different initial conditions shown in (b)-(c), respectively. All simulation results correspond to $\Delta = \tau = P_m = 1$ and $P_p = 0.001$, and the numerical solution of the continuum model is obtained with $\delta t = 1 \times 10^{-3}$ and $\epsilon = 1 \times 10^{-5}$.

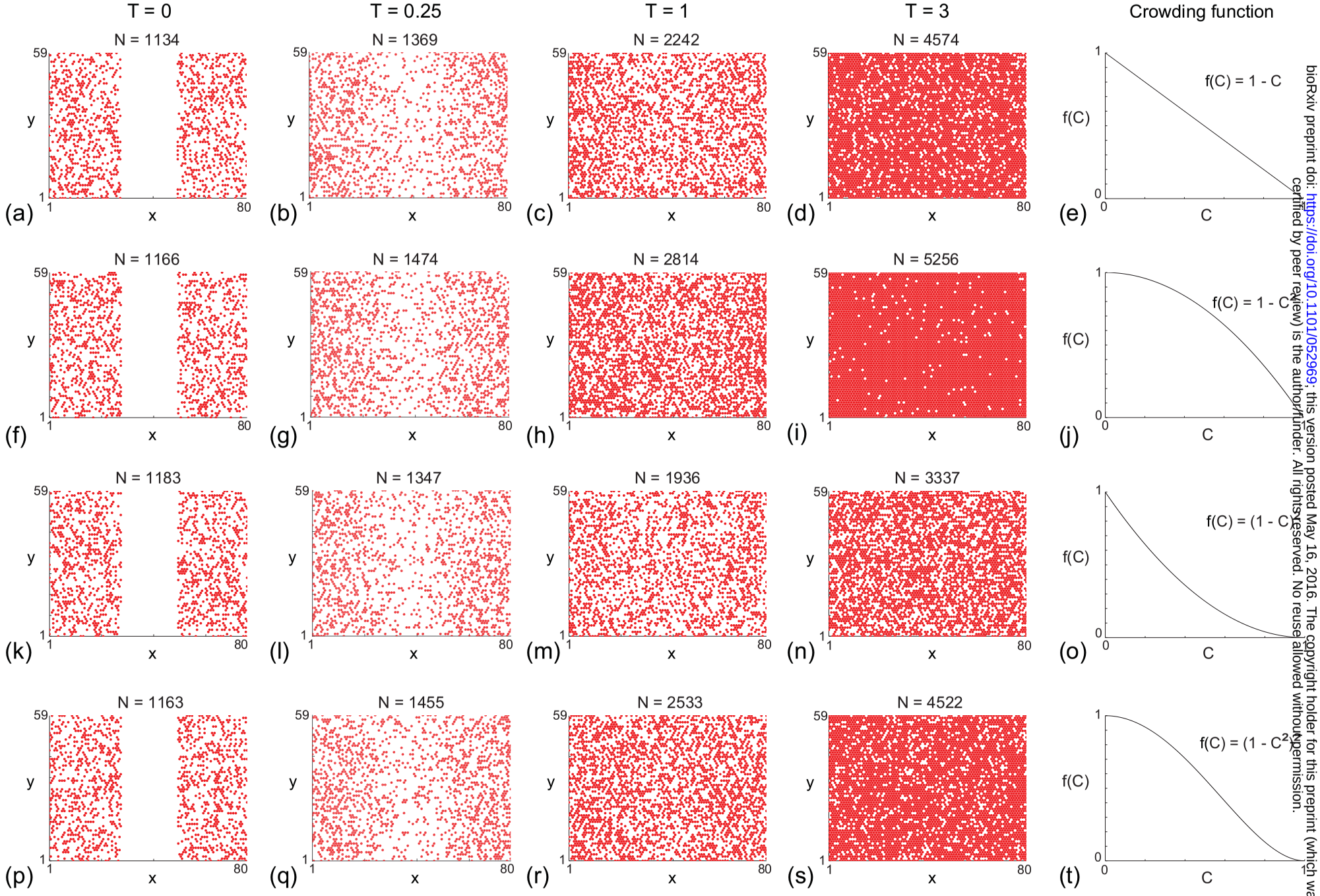

Figure 8. Snapshots of simulations for a suite of scratch assays. In each row the distributions of agents at time $\lambda t = T = 0, 0.25, 1, 3$ are shown along with the corresponding form of $f(C)$. Each simulation is initiated by randomly populating a lattice, corresponding to lattice of size $I = 80$ and $J = 68$, so that each site is occupied with probability 0.30. A scratch of 23 lattice sites wide is made at $T = 0$. All simulations correspond to $\Delta = \tau = P_m = 1$, $P_p = 0.001$ and $r = 4$.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

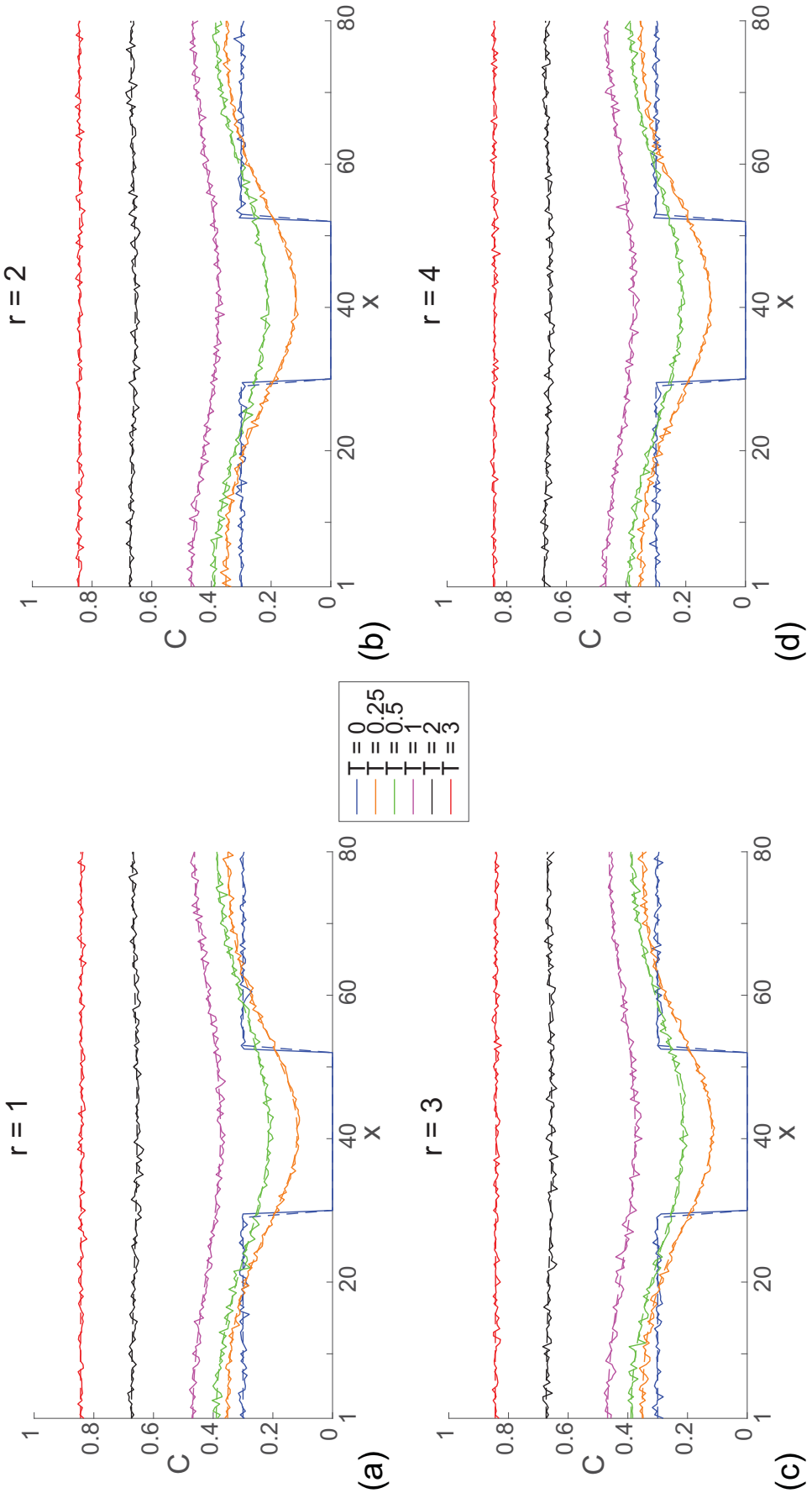

Figure 9. Comparison of averaged simulation data and the solution of the corresponding continuum model for a scratch assay with $f(C) = 1 - C$. Results in (a)-(d) compare averaged simulation data (solid lines) and the solution of the corresponding continuum model (dashed lines) for $r = 1, 2, 3$ and 4, respectively. In each subfigure, agent density profiles are given at $\lambda t = T = 0, 0.25, 0.5, 1, 2, 3$ to illustrate the temporal progression of the wound healing process. All simulation results are averaged across 100 identically prepared realisations of the discrete model, with $\Delta = \tau = P_m = 1$ and $P_p = 0.001$, on a lattice of size $I = 80$ and $J = 68$. The numerical solution of the continuum model is with $\delta x = 0.25$, $\delta t = 0.1$ and $\epsilon = 1 \times 10^{-5}$.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

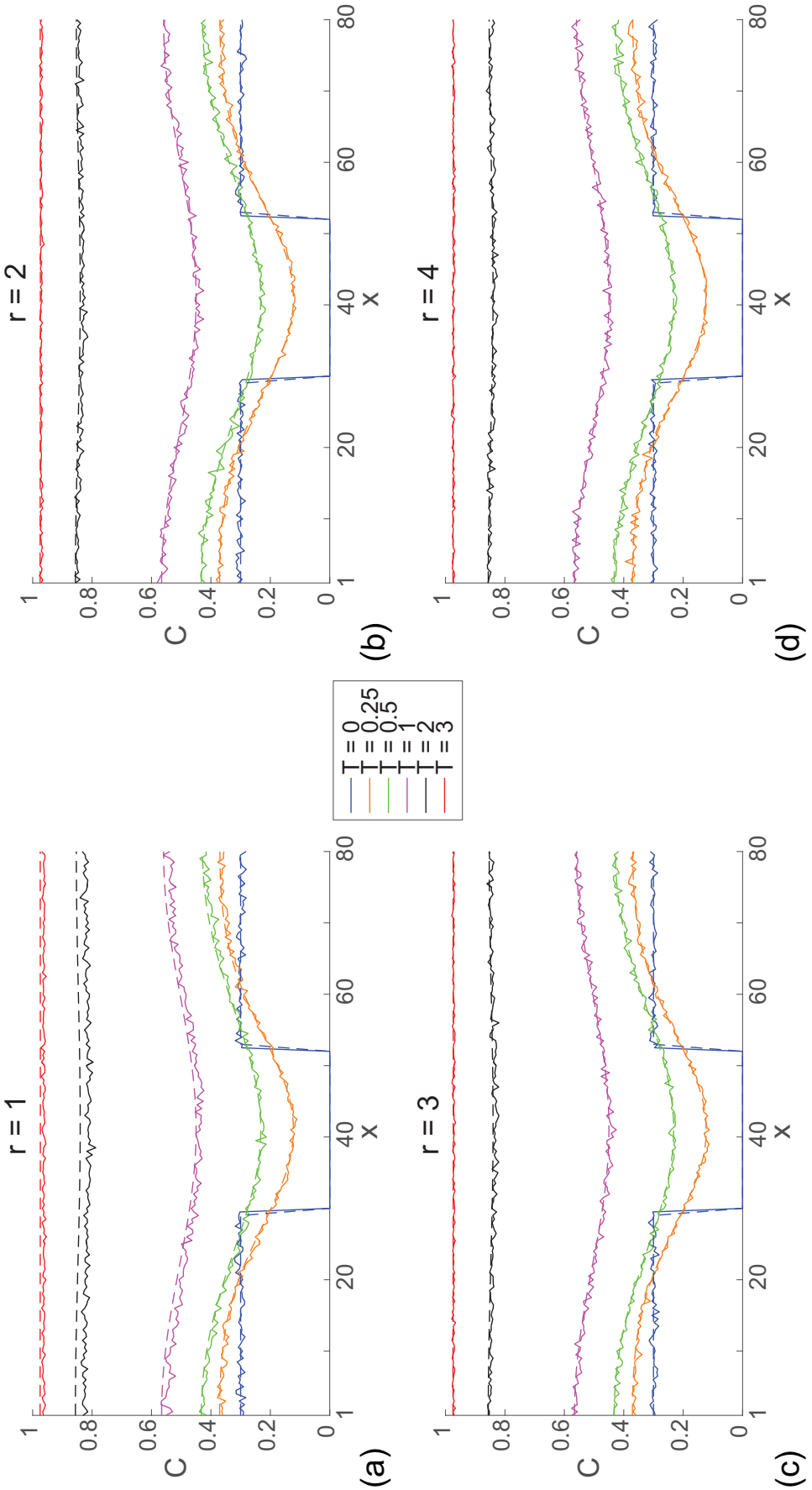

Figure 10. Comparison of averaged simulation data and the solution of the corresponding continuum model for a scratch assay with $f(C) = 1 - C^2$. Results in (a)-(d) compare averaged simulation data (solid lines) and the solution of the corresponding continuum model (dashed lines) for $r = 1, 2, 3$ and 4, respectively. In each subfigure, agent density profiles are given at $\lambda t = T = 0, 0.25, 0.5, 1, 2, 3$ to illustrate the temporal progression of the wound healing process. All simulation results are averaged across 100 identically prepared realisations of the discrete model, with $\Delta = \tau = P_m = 1$ and $P_p = 0.001$, on a lattice of size $I = 80$ and $J = 68$. The numerical solution of the continuum model is with $\delta x = 0.25$, $\delta t = 0.1$ and $\epsilon = 1 \times 10^{-5}$.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

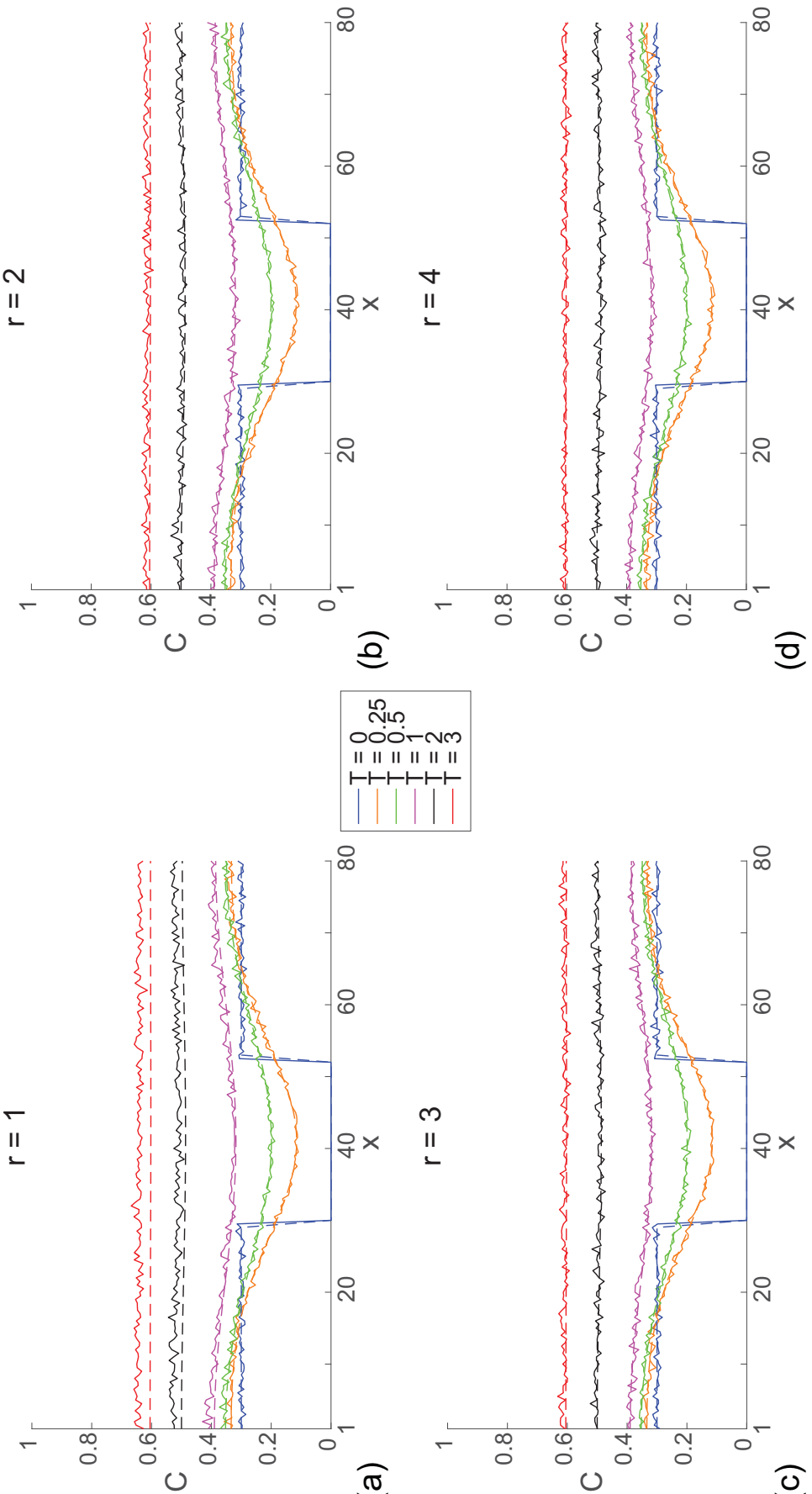

Figure 11. Comparison of averaged simulation data and the solution of the corresponding continuum model for a scratch assay with $f(C) = (1-C)^2$. Results in (a)-(d) compare averaged simulation data (solid lines) and the solution of the corresponding continuum model (dashed lines) for $r = 1, 2, 3$ and 4, respectively. In each subfigure, agent density profiles are given at $\lambda t = T = 0, 0.25, 0.5, 1, 2, 3$ to illustrate the temporal progression of the wound healing process. All simulation results are averaged across 100 identically prepared realisations of the discrete model, with $\Delta = \tau = P_m = 1$ and $P_p = 0.001$, on a lattice of size $I = 80$ and $J = 68$. The numerical solution of the continuum model is with $\delta x = 0.25$, $\delta t = 0.1$ and $\epsilon = 1 \times 10^{-5}$.

bioRxiv preprint doi: https://doi.org/10.1101/052969; this version posted May 16, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

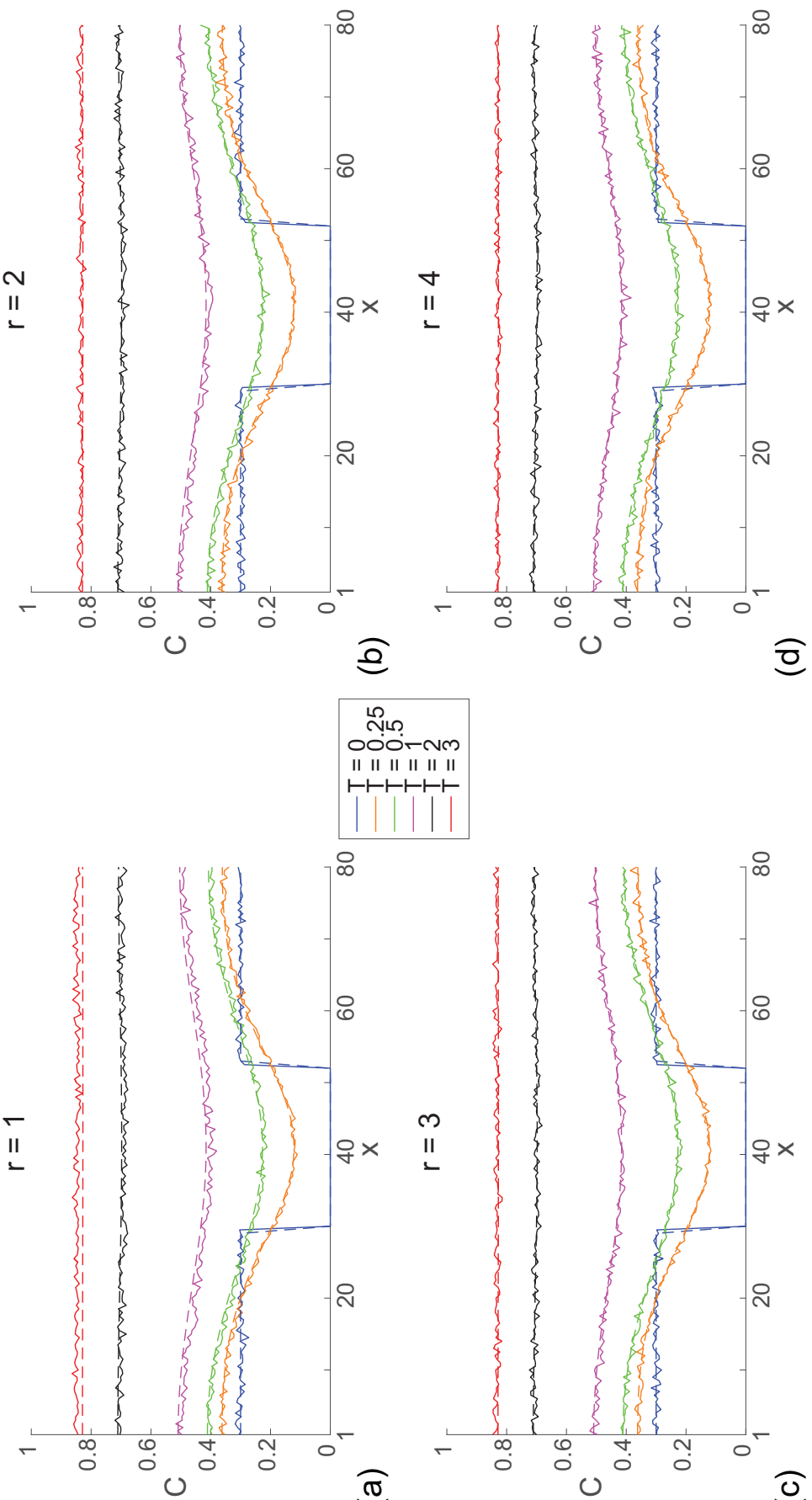

Figure 12. Comparison of averaged simulation data and the solution of the corresponding continuum model for a scratch assay with $f(C) = (1 - C^2)^2$. Results in (a)-(d) compare averaged simulation data (solid lines) and the solution of the corresponding continuum model (dashed lines) for $r = 1, 2, 3$ and 4, respectively. In each subfigure, agent density profiles are given at $\lambda t = T = 0, 0.25, 0.5, 1, 2, 3$ to illustrate the temporal progression of the wound healing process. All simulation results are averaged across 100 identically prepared realisations of the discrete model, with $\Delta = \tau = P_m = 1$ and $P_p = 0.001$, on a lattice of size $I = 80$ and $J = 68$. The numerical solution of the continuum model is with $\delta x = 0.25$, $\delta t = 0.1$ and $\epsilon = 1 \times 10^{-5}$.