Bulletin of Mathematical Biology (2025) 87:77
https://doi.org/10.1007/s11538-025-01452-y

Society for Mathematical Biology

ORIGINAL ARTICLE


# Spatial Segregation Across Travelling Fronts in Individual-Based and Continuum Models for the Growth of Heterogeneous Cell Populations

**José A. Carrillo**[1] · **Tommaso Lorenzi**[2] · **Fiona R. Macfarlane**[3,4]

Received: 12 December 2024 / Accepted: 15 April 2025


## Abstract

We consider a partial differential equation model for the growth of heterogeneous cell populations subdivided into multiple distinct discrete phenotypes. In this model, cells preferentially move towards regions where they are less compressed, and thus their movement occurs down the gradient of the cellular pressure. The cellular pressure is defined as a weighted sum of the densities (i.e. the volume fractions) of cells with different phenotypes. To translate into mathematical terms the idea that cells with distinct phenotypes have different morphological and mechanical properties, both the cell mobility and the weighted amount the cells contribute to the cellular pressure vary with their phenotype. We formally derive this model as the continuum limit of an on-lattice individual-based model, where cells are represented as single agents undergoing a branching biased random walk corresponding to phenotype-dependent and pressure-regulated cell division, death, and movement. Then, we study travelling wave solutions whereby cells with different phenotypes are spatially segregated across the invading front. Finally, we report on numerical simulations of the two models, demonstrating excellent agreement between them and the travelling wave analysis. The results presented here indicate that inter-cellular variability in mobility can support the maintenance of spatial segregation across invading fronts, whereby cells with a higher mobility drive invasion by occupying regions closer to the front edge.

**Keywords** Phenotypic heterogeneity · Individual-based models · Continuum models · Travelling waves · Spatial segregation

✉ José A. Carrillo
carrillo@maths.ox.ac.uk

1 Mathematical Institute, University of Oxford, Oxford, UK

2 Department of Mathematical Sciences "G. L. Lagrange", Politecnico di Torino, Turin, Italy

3 School of Mathematics and Statistics, University of St Andrews, St Andrews, UK

4 Applied BioSimulation, Certara (UK), Sheffield, UK

Published online: 19 May 2025

# 1 Introduction

## 1.1 Background

Systems of partial differential equations (PDEs) modelling the growth of populations that disperse to avoid crowding (Bertsch et al. 1985; Gurtin and Pipkin 1984) have been applied to describe the spatiotemporal dynamics of multiple types of cells underpinning tissue development, wound healing, and tumour growth (Ambrosi and Preziosi 2002; Bertsch et al. 2012; Bubba et al. 2020; Byrne and Preziosi 2003; Carrillo 2019; Chaplain 2006; Ciarletta et al. 2011; David et al. 2024; Drasdo and Hoehme 2012; Giverso et al. 2022; Gwiazda et al. 2019; Lorenzi et al. 2017; Bertsch et al. 2010; Oelschläger 1989; Roose et al. 2007; Preziosi and Tosin 2009).

Focusing on a one-dimensional spatial scenario, and considering two cell types, a prototypical example of these models is provided by the following PDE system

$$\begin{cases} \partial_t n_1 - \partial_x \left(n_1 \, \partial_x \rho\right) = G_1(\rho)\, n_1, & \text{(1a)} \\ \partial_t n_2 - \partial_x \left(n_2 \, \partial_x \rho\right) = G_2(\rho)\, n_2, \quad (t,x) \in (0,\infty) \times \mathbb{R} & \text{(1b)} \\ \rho(t,x) := n_1(t,x) + n_2(t,x), & \text{(1c)} \end{cases}$$

Here, the functions $n_1(t,x)$ and $n_2(t,x)$ are the densities (i.e. the volume fractions) of cells of types 1 and 2 at position $x$ at time $t$, and the function $\rho(t,x)$ defined via (1c) is the total cell density (i.e. the total cell volume fraction). The transport terms on the left-hand sides of the PDEs (1a) and (1b) model the effect of cell movement and capture the tendency of cells to move away from overcrowded regions (i.e. to move down the gradient of the total cell density) (Chaplain 2006). Moreover, the reaction terms on the right-hand sides model the effect of cell division and death, with the functions $G_1(\rho)$ and $G_2(\rho)$ being the net growth rates of the densities of cells of types 1 and 2. These functions depend on the total cell density so as to integrate the effect of density-dependent inhibition of growth (i.e. the cessation of growth at sufficiently high cell density) (Lieberman 1981).

A variation on the model (1) is given by the following PDE system

$$\begin{cases} \partial_t n_1 - \partial_x \left(n_1 \, \partial_x p\right) = G_1(p)\, n_1, & \text{(2a)} \\ \partial_t n_2 - \partial_x \left(n_2 \, \partial_x p\right) = G_2(p)\, n_2, \quad (t,x) \in (0,\infty) \times \mathbb{R} & \text{(2b)} \\ p(t,x) := \Pi[\rho](t,x), \rho(t,x) := n_1(t,x) + n_2(t,x), & \text{(2c)} \end{cases}$$

The PDE system (2) can be obtained by replacing the cell density $\rho(t,x)$ in (1) with the cellular pressure $p(t,x)$ and then closing the resulting system for the cell density functions by defining the cell pressure as a function of the total cell density through an appropriate constitutive relation $\Pi[\rho]$ (Ambrosi and Preziosi 2002; Perthame et al. 2014). This makes it possible to incorporate into the model the effects of pressure-regulated growth – i.e. the fact that cells will stop dividing if the pressure at their current position overcomes a critical value (Byrne and Preziosi 2003; Byrne and Drasdo 2009; Drasdo and Hoehme 2012; Ranft et al. 2010) – and pressure-regulated cell movement – i.e. the movement of cells down the gradient of the cellular pressure towards regions where they are less compressed (Byrne and Chaplain 1997; Byrne and Preziosi 2003).

These models, which provide a population-level description of cell dynamics, can be derived as the continuum limits of underlying individual-based models, which track the dynamics of single cells and are thus able to capture the finer details of single-cell movement, division, and death (Drasdo 2005). To this end, a range of limiting procedures have been developed and employed for transitioning between individual-based models for the growth of cell populations and continuum models formulated as systems of PDEs of the form of (1) and (2) and related forms (Alasio et al. 2022; Carrillo 2019; Chaplain et al. 2020; Dyson et al. 2012; Fozard et al. 2010; Galiano and Selgas 2014; Lorenzi et al. 2020; Penington et al. 2011; Pillay et al. 2017, 2018; Simpson et al. 2011, 2024).

An interesting feature of PDE systems like (1) and (2) is that they can support spatial segregation between different cell types – i.e. the fact that if cells of different types are initially separated (i.e. they occupy distinct regions of the spatial domain at the initial time) then they will remain separated also at later times. Mathematically speaking, this means that the solutions of these PDE systems may exhibit a propagation of segregation property (David et al. 2024). This is an interesting mathematical aspect, the study of which has received increasing attention in recent decades (Bertsch et al. 2015, 2012, 1987a, b, 1985; Burger et al. 2020; Carrillo 2018a, b; Girardin and Hilhorst 2022; Galiano and Selgas 2014; Lorenzi et al. 2017; Bertsch et al. 2010; Falcó et al. 2024). Furthermore, it makes such models an appropriate theoretical framework to investigate the mechanisms underlying cell segregation processes leading to the formation of sharp borders between cells of distinct types or with different phenotypes, which is observed both in normal development and in tumourigenesis (Batlle and Wilkinson 2012).

### 1.2 Object of Study

In this paper, we consider the following model for the growth of a phenotypically heterogeneous population comprising cells of different types (i.e. with different phenotypes), which are labelled by the index $i = 1, \ldots, I$ with $I \geq 2$:

$$\begin{cases} \partial_t n_1 - \mu_1 \, \partial_x \left( n_1 \, \partial_x p \right) = G_1(p) \, n_1 & \text{(3a)} \\ \partial_t n_i - \mu_i \, \partial_x \left( n_i \, \partial_x p \right) = 0, \quad i = 2, \ldots, I \qquad (t, x) \in (0, \infty) \times \mathbb{R} & \text{(3b)} \\ p(t, x) := \displaystyle\sum_{i=1}^{I} \omega_i \, n_i(t, x), & \text{(3c)} \end{cases}$$

In the model (3), the functions $n_1(t, x), \ldots, n_I(t, x)$ represent the densities (i.e. the volume fractions) of cells with phenotypes $1, \ldots, I$ at position $x$ at time $t$, while the function $p(t, x)$ represents the cellular pressure, which is defined as a function of the cell densities through the constitutive relation (3c). The positive parameters $\omega_1, \ldots, \omega_I$ provide a measure of the weighted amounts that cells with phenotypes $1, \ldots, I$ contribute towards the cellular pressure, and the values of these parameters are related to the morphological and mechanical properties of the cells (such as cell size and stiffness), which may vary depending on the cell phenotype (Masaeli et al. 2016). In analogy with model (2), the second terms on the left-hand sides of the PDEs (3a)

and (3b) are the rates of change of the densities of cells with phenotypes $1, \dots, I$ due to pressure-regulated cell movement. The positive parameter $\mu_i$ is the mobility coefficient of cells with phenotype $i$ (Ambrosi and Preziosi 2002; Byrne and Drasdo 2009), which in model (2) is implicitly assumed to be the same for all cells (and it is then set to 1), but in fact it can vary due again to differences in morphological and mechanical properties (such as cell elongation and nucleus deformability) between cells with different phenotypes (Kalukula et al. 2022; Lamouille et al. 2014). Moreover, similarly to model (2), the term on the right-hand side of the PDE (3a) is the rate of change of the density of cells with phenotype 1 due to pressure-regulated growth and, therefore, the function $G_1(p)$ is the net growth rate of the density of cells with phenotype 1 when exposed to the cellular pressure $p$. Focusing on the impact of inter-cellular variability in mobility on cell dynamics rather than variability in division and death rates, the right-hand sides of the PDEs (3b) are set to zero. This corresponds to a biological scenario where division and death of cells with phenotypes $2, \dots, I$ occur on much slower time scales compared to division and death of cells with phenotype 1, and can thus be neglected. For example, in the context of cancer, tumour-derived leader cells undergo epithelial to mesenchymal transition acquiring mesenchymal characteristics including significantly increased motility and significantly reduced proliferation (Chen et al. 2020; Huang et al. 2022; Konen et al. 2017; Vilchez Mercedes et al. 2021; Wang et al. 2023; Zanotelli et al. 2021). Specifically, it has been observed that on short time-scales the proliferation rate of leader cells compared to follower cells is negligible (Konen et al. 2017). Hence, in the framework of model (3), cells with phenotype 1 could be regarded as follower-type cells, while cells with phenotypes $2, \dots, I$ could be regarded as leader-type cells with heterogeneous morphological and mechanical properties.

In contrast to previous works on related models, which considered two cell phenotypes only (i.e. $i = 1, 2$) and assumed the values of the weights $\omega_i$ and the values of the mobility coefficients $\mu_i$ to be the same for both phenotypes, in this paper we consider an arbitrary number $I \geq 2$ of cell phenotypes and allow both the values of $\omega_i$ and the values of $\mu_i$ to vary with the phenotype, in order to capture intra-population phenotypic heterogeneity more accurately.

### 1.3 Outline of the Paper

Building on the modelling framework developed in Chaplain et al. (2020), we first formulate an individual-based model (see Sect. 2) where cells are represented as single agents undergoing phenotype-dependent and pressure-regulated cell division, death, and movement according to a set of rules. These rules result in cells performing a branching biased random walk over the one-dimensional lattice that represents the spatial domain (Hughes 1996; Johnston et al. 2012; Penington et al. 2011). Then, using a limiting procedure analogous to the one that we employed in Bubba et al. (2020); Macfarlane and Chaplain (2020); Macfarlane (2022), we formally derive model (3) as the continuum limit of this individual-based model (see Sect. 2.2 and Appendix A:). After that, generalising the method of proof that developed in Lorenzi et al. (2017), we carry out travelling wave analysis of the model (3) (see Sect. 3) and study, under appropriate

assumptions on the function $G_1$ and the mobility coefficients $\mu_1, \ldots, \mu_I$, travelling front solutions wherein cells with phenotypes labelled by different values of the index $i = 1, \ldots, I$ are spatially segregated across the front (i.e. they occupy distinct regions of the front). Finally, we report on numerical simulations of the individual-based model and numerical solutions of the PDE model (3), which demonstrate excellent agreement between the two models, thus validating the formal limiting procedure employed to derive the continuum limit of the individual-based model, and confirm the results of travelling wave analysis (see Sect. 4). We conclude with a discussion of the main results obtained and provide a brief overview of possible research perspectives (see Sect. 5).

## 2 Individual-Based and Continuum Models

### 2.1 Individual-Based Model

Considering a one-dimensional spatial scenario, we let cells be distributed and move along the real line $\mathbb{R}$. We introduce the notation $\mathbb{R}_+ := \{z \in \mathbb{R} : z \geq 0\}$, $\mathbb{R}_+^* := \mathbb{R}_+ \setminus \{0\}$, and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, and then discretise the time variable $t \in \mathbb{R}_+$ and the space variable $x \in \mathbb{R}$ as $t_k = k\tau$ with $k \in \mathbb{N}_0$ and $x_j = j\Delta_x$ with $j \in \mathbb{Z}$, respectively, where $\tau \in \mathbb{R}_+^*$ represents the time-step and $\Delta_x \in \mathbb{R}_+^*$ represents the space-step.

We consider a population comprising cells expressing one amongst $I \geq 2$ distinct discrete phenotypes, each labelled by an index $i = 1, \ldots, I$, and define the density of cells with phenotype $i$ at position $x_j$ at time $t_k$, denoted $n_{i,j}^k$, as

$$n_{i,j}^k := \frac{N_{i,j}^k}{\Delta_x}, \tag{4}$$

where $N_{i,j}^k$ is the number of cells with phenotype $i$ at position $x_j$ at time $t_k$. Moreover, we define the cellular pressure at position $x_j$ at time $t_k$, denoted $p_j^k$, as a function of $n_{i,j}^k$ through the following constitutive relation:

$$p_j^k := \sum_{i=1}^{I} \omega_i \, n_{i,j}^k, \quad \text{with} \quad \omega_i \in \mathbb{R}_+^*. \tag{5}$$

As mentioned in Sect. 1, the constitutive relation (5) translates into mathematical terms the idea that cells with each phenotype $i$ may contribute a different weighted amount, which is represented by the parameter $\omega_i$, to the cellular pressure. The value of the parameter $\omega_i$ is related to the morphological and mechanical properties of the cells; for instance, higher values of $\omega_i$ may correspond to larger cell size and/or larger cell stiffness (Masaeli et al. 2016). The dynamics of the cells are governed by the rules summarised in Fig. 1 and detailed in the following subsections.

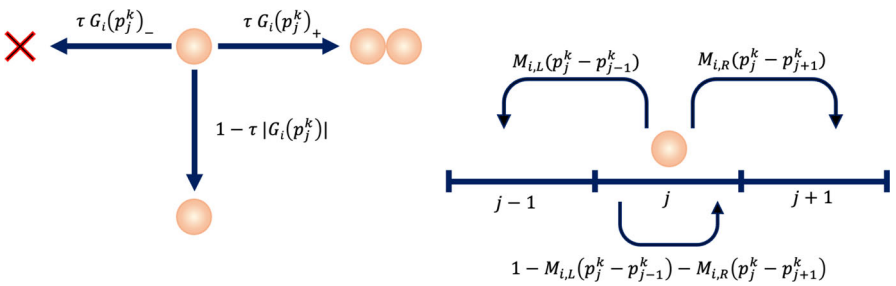


**Fig. 1** Schematic overview of the mechanisms incorporated in the individual-based model. Between time-steps $k$ and $k+1$, each cell with phenotype $i = 1, \ldots, I$ at spatial position $x_j$ may: divide with probability $\tau\, G_i(p_j^k)_+$, die with probability $\tau\, G_i(p_j^k)_-$, and remain quiescent with probability $1 - \tau\left(G_i(p_j^k)_+ + G_i(p_j^k)_-\right) = 1 - \tau\,|G_i(p_j^k)|$ (left panel); move to spatial positions $x_{j-1}$ or $x_{j+1}$ with probabilities $M_{i,L}(p_j^k - p_{j-1}^k)$ or $M_{i,R}(p_j^k - p_{j+1}^k)$, respectively, or remain stationary with probability $1 - M_{i,L}(p_j^k - p_{j-1}^k) - M_{i,R}(p_j^k - p_{j+1}^k)$ (right panel)

### 2.1.1 Modelling Cell Division and Death

We model pressure-regulated cell division and death as a branching process along the spatial dimension, whereby cells divide and die with probabilities that depend on both their phenotype and the pressure they experience, as illustrated by the schematic in the left panel of Fig. 1. If cell division occurs, a dividing cell is instantly replaced by two identical progeny cells that inherit the spatial position and phenotype of the parent cell. Conversely, a cell undergoing cell death is instantly removed from the population.

We introduce the function $G_i(p_j^k)$, which represents the net growth rate of the density of cells with phenotype $i = 1, \ldots, I$ at spatial position $x_j$ at time $t_k$, and use this function to define the probability of cell division and death in the individual-based model. Specifically, between time-steps $k$ and $k+1$ we let a cell with phenotype $i$ at position $x_j$ divide with probability

$$\tau G_i(p_j^k)_+, \quad \text{where} \quad (\cdot)_+ = \max(0, \cdot),$$

die with probability

$$\tau G_i(p_j^k)_-, \quad \text{where} \quad (\cdot)_- = -\min(0, \cdot),$$

and remain quiescent with probability

$$1 - \tau\left(G_i(p_j^k)_+ + G_i(p_j^k)_-\right) = 1 - \tau|G_i(p_j^k)|.$$

By choosing the time-step $\tau$ sufficiently small, we ensure that all the quantities above are between 0 and 1.

In order to take into account the fact that cells will stop dividing if the pressure at their current position exceeds a critical value, known as the homeostatic pressure

(Basan et al. 2009), which is modelled by the parameter $\overline{p} \in \mathbb{R}_+^*$, and considering a scenario in which cells with phenotypes labelled by different values of the index $i$ may undergo cell division and death over different time scales, we make the following assumptions:

$$G_i(p) := \alpha_i\, G(p), \quad G(0) < \infty, \quad G(\overline{p}) = 0, \quad \frac{\mathrm{d}G}{\mathrm{d}p} < 0 \,\forall p \in \mathbb{R}_+. \tag{6}$$

Here the parameter $\alpha_i \in \mathbb{R}_+$ is linked to the time scale over which cells with phenotype $i$ undergo cell division and death. In particular, focusing on the case where division and death of cells with phenotypes $i = 2, \ldots, I$ can be neglected, since they occur on much slower time scales compared to division and death of cells with phenotype $i = 1$, we also assume

$$\alpha_1 > 0 \quad \text{and} \quad \alpha_i = 0 \;\text{ for }\; i = 2, \ldots, I. \tag{7}$$

### 2.1.2 Modelling Cell Movement

We model directional cell movement in response to pressure differentials as a biased random walk, whereby the movement probabilities depend on the difference between the cellular pressure at the position occupied by the cell and the cellular pressure at neighbouring positions, as illustrated by the schematic in the right panel of Fig. 1. In particular, we assume that cells move down the gradient of the pressure so as to reach regions where they are less compressed. Moreover, in order to capture the fact that the phenotype of the cells determines their sensitivity to the pressure gradient, and thus their mobility, we introduce the parameters $\gamma_i \in \mathbb{R}_+^*$ to model the sensitivity to the pressure gradient of cells with phenotypes $i = 1, \ldots, I$.

In the individual-based model, between time-steps $k$ and $k+1$ we let a cell with phenotype $i$ at position $x_j$ move to position $x_{j-1}$ (i.e. move left) with probability

$$M_{i,L}(p_j^k - p_{j-1}^k),$$

move to position $x_{j+1}$ (i.e. move right) with probability

$$M_{i,R}(p_j^k - p_{j+1}^k),$$

and remain stationary with probability

$$1 - \left(M_{i,L}(p_j^k - p_{j-1}^k) + M_{i,R}(p_j^k - p_{j+1}^k)\right).$$

Specifically, recalling that, as described in Sect. 2.1.1, the parameter $\overline{p}$ represents the homeostatic pressure, we use the following definitions

$$M_{i,L}(p_j^k - p_{j-1}^k) = \gamma_i \frac{\left(p_j^k - p_{j-1}^k\right)_+}{2\overline{p}},$$
$$M_{i,R}(p_j^k - p_{j+1}^k) = \gamma_i \frac{\left(p_j^k - p_{j+1}^k\right)_+}{2\overline{p}}, \quad \text{where} \quad (\cdot)_+ = \max(0, \cdot), \tag{8}$$

and choose the model parameters and functions such that $0 \leq M_{i,L}(p_j^k - p_{j-1}^k) + M_{i,R}(p_j^k - p_{j+1}^k) < 1$ for all $i$, $j$, and $k$.

Without loss of generality, considering a scenario where cells with phenotypes labelled by smaller values of the index $i$ have a lower sensitivity to the pressure gradient, and thus a lower mobility, we assume

$$0 < \gamma_1 < \gamma_2 < \ldots < \gamma_{I-1} < \gamma_I. \tag{9}$$

***Remark 1*** Taken together, assumptions (6)-(7) and (9) correspond to the situation in which, due to a trade-off between cell proliferative and migratory abilities, fast-dividing cells with phenotype $i = 1$ display the lowest mobility (Chen et al. 2020; Huang et al. 2022; Konen et al. 2017; Vilchez Mercedes et al. 2021; Wang et al. 2023; Zanotelli et al. 2021).

### 2.2 Corresponding Continuum Model

As detailed in Appendix A:, from the branching biased random walk underlying the individual-based model presented in the previous section, one formally derives, as the corresponding continuum limit, a PDE system for the functions $n_i(t, x)$, each modelling the density (i.e. the volume fraction) of cells with phenotype $i = 1, \ldots, I$, at position $x \in \mathbb{R}$ at time $t \in \mathbb{R}_+$. This is done through an extension of the limiting procedure that we employed in Bubba et al. (2020); Macfarlane and Chaplain (2020); Macfarlane (2022) – see also Simpson et al. (2007); Johnston et al. (2012); Simpson et al. (2011, 2024) for related strategies.

In summary, one writes down a balance equation for the density of cells with phenotype $i$ at spatial position $x_j$ at time $t_{k+1}$, which depends on cell densities at time $t_k$ at position $x_j$ and neighbouring positions $x_{j-1}$ and $x_{j+1}$, as a result of cell movement and cell division and death. Specifically, one has

$$n_{i,\,j}^{k+1} = n_{i,\,j}^k \left\{ \left(1 + \tau G_i(p_j^k)\right) \left[ 1 - \frac{\gamma_i \left(p_j^k - p_{j+1}^k\right)_+}{2\overline{p}} - \frac{\gamma_i \left(p_j^k - p_{j-1}^k\right)_+}{2\overline{p}} \right] \right\}$$
$$+ n_{i,\,j+1}^k \left\{ \left(1 + \tau G_i(p_{j+1}^k)\right) \left[ \frac{\gamma_i \left(p_{j+1}^k - p_j^k\right)_+}{2\overline{p}} \right] \right\}$$

$$+n^k_{i,\,j-1}\left\{\left(1+\tau G_i(p^k_{j-1})\right)\left[\frac{\gamma_i\left(p^k_{j-1}-p^k_j\right)_+}{2\overline{p}}\right]\right\}. \tag{10}$$

From the balance equation (10), employing a formal limiting procedure that includes letting $\Delta_x \to 0$ and $\tau \to 0$ in such a way that

$$\frac{\gamma_i \Delta_x^2}{2\tau\overline{p}} \to \mu_i, \quad \text{where} \quad \mu_i \in \mathbb{R}^*_+,\ i=1,\ldots,I, \tag{11}$$

under the constitutive relation (5) one formally obtains the PDE system

$$\begin{cases} \partial_t n_i - \mu_i\,\partial_x\left(n_i\,\partial_x p\right) = G_i(p)\,n_i, \quad i=1,\ldots,I, \\ \\ p(t,x) := \displaystyle\sum_{i=1}^{I}\omega_i\,n_i(t,x), \end{cases} \qquad (t,x)\in(0,\infty)\times\mathbb{R}, \tag{12}$$

subject to the following assumptions on the mobility coefficients, $\mu_i$,

$$0<\mu_1<\mu_2<\ldots<\mu_{I-1}<\mu_I. \tag{13}$$

Assumptions (13) descend from assumptions (9) when conditions (11) hold. Under the additional assumptions (6)-(7) the PDE system (12) then reduces to the PDE system (3).

## 3 Travelling Wave Analysis

In this section, under assumptions (6)-(7) and (13), we carry out travelling wave analysis of the continuum model (3). Substituting the travelling wave ansatz

$$n_i(t,x)=n_i(z), \quad z:=x-c\,t, \quad c\in\mathbb{R}^*_+,$$

where $c$ is the travelling wave speed, into the model (3) yields

$$\begin{cases} -c\,n_1' - \mu_1\left(n_1\,p'\right)' = G_1(p)\,n_1, & \text{(14a)} \\ -c\,n_i' - \mu_i\left(n_i\,p'\right)' = 0, \quad i=2,\ldots,I, \quad z\in\mathbb{R}. & \text{(14b)} \\ p(z) := \displaystyle\sum_{i=1}^{I}\omega_i\,n_i(z), & \text{(14c)} \end{cases}$$

We seek travelling wave solutions that satisfy the following conditions

$$n_i(z)\begin{cases} >0, & \text{for } z\in(z_{i-1},z_i) \\ =0, & \text{for } z\notin(z_{i-1},z_i), \end{cases} \qquad i=1,\ldots,I, \tag{15}$$

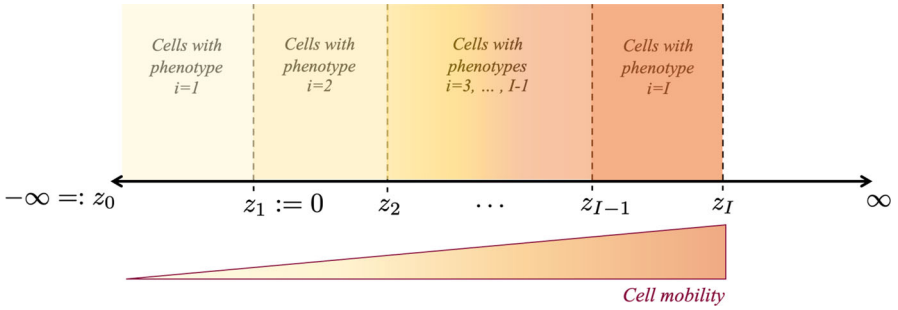


**Fig. 2** Schematic overview of spatial segregation across travelling waves. Schematic of how, under assumptions (6)-(7) and (13), cells with phenotypes labelled by different values of the index $i$ are spatially segregated across invading fronts, which are represented by travelling wave solutions of the continuum model (3), i.e. solutions of the system of differential equations (14), subject to conditions (15)-(16)

where

$$-\infty =: z_0 < z_1 := 0 < z_2 < \ldots < z_{I-1} < z_I < \infty, \tag{16}$$

along with the asymptotic (i.e. boundary) condition

$$p(-\infty) = \overline{p}. \tag{17}$$

Note that, under conditions (15)-(16), conservation of mass ensures that

$$\int_{z_{i-1}}^{z_i} n_i(z)\,\mathrm{d}z = M_i, \quad M_i \in \mathbb{R}_+^*, \quad i = 2, \ldots, I, \tag{18}$$

where the parameter $M_i$ represents the number (i.e. the total volume fraction) of cells with phenotype $i = 2, \ldots, I$ in the population.

***Remark 2*** Conditions (15)-(16) correspond to a scenario in which cells with phenotypes labelled by different values of the index $i$ are spatially segregated across invading fronts, which are represented by travelling wave solutions of the continuum model (3), i.e. solutions of the system of differential equations (14). More precisely, also in the light of assumptions (6)-(7) and (13), cells with phenotype $i = 1$ (i.e. fast-dividing cells with the lowest mobility) make up the bulk of the population in the rear of the wave, while cells with phenotypes labelled by increasing values of $i > 1$, which display a higher mobility, are found in the regions closer to the invading edge (cf. the schematic in Fig. 2).

The properties of such travelling wave solutions are established by Theorem 1.

**Theorem 1** *Let assumptions (6)-(7) and (13) hold. For any $M_2, \ldots, M_I \in \mathbb{R}_+^*$ there exist $z_2, \ldots, z_I \in \mathbb{R}_+^*$ and $c \in \mathbb{R}_+^*$ such that the system of differential equations (14) subject to conditions (15)-(17) admits a solution wherein each component $n_i(z)$ is positive, continuous, and decreasing on $(z_{i-1}, z_i)$ for $i = 1, \ldots, I$, and the components $n_2(z), \ldots, n_I(z)$ satisfy conditions (18) as well. Moreover, the function $p(z)$*

*defined via the constitutive relation (14c) is positive, continuous, and decreasing on* $(-\infty, z_I)$*, with*

$$p(0) = \sqrt{2\,c \sum_{j=2}^{I} \frac{\omega_j}{\mu_j} M_j}, \tag{19}$$

*and it has a kink at the points* $0, z_2, \ldots, z_{I-2}, z_{I-1}$ *with*

$$\begin{aligned} &-\mu_2\, p'(0^+) = -\mu_1\, p'(0^-) = c, \\ &-\mu_{i+1}\, p'(z_i^+) = -\mu_i\, p'(z_i^-) = c, \;\; i = 2, \ldots, I-1. \end{aligned} \tag{20}$$

***Proof*** We prove Theorem 1 in 8 steps. In summary, building on the shooting method that we employed in Lorenzi et al. (2017), first we prove, for $c \in \mathbb{R}_+^*$ given, that: (i) The cell pressure $p(z)$ and the cell density $n_i(z)$ are positive, continuous, and monotonically decreasing on the intervals $(z_{i-1}, z_i)$ for all $i = 1, \ldots, I$. Then $p(z)$ is not only monotonically decreasing on each interval but as a whole in $\mathbb{R}$, with possibly jumps at the points $z_i$ for $i = 1, \ldots, I-1$. This along with the non-negativity of $p(z)$ ensures that $0 \leq p(z) \leq \overline{p}$ for all $z \in \mathbb{R}$ (*Step 1*). (ii) The cell pressure is continuous also in $z_i$ for all $i = 1, \ldots, I$ (*Step 2*). (iii) The derivative of the cell pressure satisfies conditions (20) (*Steps 3-4*). Then, still for $c \in \mathbb{R}_+^*$ given, imposing conditions (18), we find the values attained by the cell pressure $p(z)$ at the points $z_2, \ldots, z_{I-1}$, and we prove that condition (19) holds (*Steps 5-6*). Next, still for $c \in \mathbb{R}_+^*$ given, we identify the points $z_2, \ldots, z_I$ (*Step 7*). Finally, we prove that there exists a unique pair $(c, p)$ that satisfies the travelling wave problem (*Step 8*).

*Preliminary observations.* Throughout the proof we will be exploiting the fact that the system of differential equations (14) subject to conditions (15)-(16) can be rewritten as

$$\begin{cases} -c\, n_1' - \mu_1 \left(n_1\, p'\right)' = G_1(p)\, n_1, \quad z \in (z_0, z_1) \equiv (-\infty, 0), & \text{(21a)} \\ -c\, n_i' - \mu_i \left(n_i\, p'\right)' = 0, \quad z \in (z_{i-1}, z_i) \quad i = 2, \ldots, I, & \text{(21b)} \\ p(z) := \omega_i\, n_i(z), \quad z \in (z_{i-1}, z_i), \quad i = 1, \ldots, I. & \text{(21c)} \end{cases}$$

Moreover, multiplying both sides of the differential equation (21a) by $\omega_1$ and both sides of the differential equation (21b) by $\omega_i$, using the relation between $p$ and $n_1$ and $p$ and $n_i$ given by (21c) along with the fact that $G_1(p) := \alpha_1\, G(p)$ where $\alpha_1 > 0$ (cf. assumptions (6)-(7)), we obtain the following set of differential equations for $p$

$$\begin{cases} -c\, p' - \mu_1 \left(p\, p'\right)' = \alpha_1\, G(p)\, p, \quad z \in (z_0, z_1) \equiv (-\infty, 0), & \text{(22a)} \\ -c\, p' - \mu_i \left(p\, p'\right)' = 0, \quad z \in (z_{i-1}, z_i) \quad i = 2, \ldots, I. & \text{(22b)} \end{cases}$$

Furthermore, we also notice that, introducing the notation

$$\mu(z) := \sum_{i=1}^{I} \mu_i\, \mathbb{1}_{(z_{i-1}, z_i)}(z), \tag{23}$$

where $\mathbb{1}_{(\cdot)}(z)$ is the indicator function of the set $(\cdot)$, and using the fact that, under conditions (15)-(16), both $p(z) = 0$ and $p'(z) = 0$ for all $z > z_I$, the set of differential equations (22b) can be rewritten in a more compact form as

$$-c\,p'(z) - \big(\mu(z)\,p(z)\,p'(z)\big)' = 0, \quad z \in (z_1, \infty) \equiv (0, \infty). \tag{24}$$

*Step 1.* For $c \in \mathbb{R}_+^*$ given, under assumptions (6) on the function $G(p)$, the differential equation (22a) subject to the condition (17) admits solutions which are positive, continuous and, by the maximum principle, monotonically decreasing on $(z_0, z_1) \equiv (-\infty, 0)$. Moreover, still for $c \in \mathbb{R}_+^*$ given, integrating the differential equation (24) between $z \in (z_{i-1}, z_i)$ with $i = 2, \ldots, I$ and $\infty$, and using the fact that both $p(z) \to 0$ and $p'(z) \to 0$ as $z \to \infty$, yields

$$p'(z) = -\frac{c}{\mu_i} \quad z \in (z_{i-1}, z_i), \quad i = 2, \ldots, I, \tag{25}$$

from which we see that the differential equation (24) admits solutions which are positive, continuous, and monotonically decreasing on $(z_{i-1}, z_i)$ for all $i = 2, \ldots, I$. In particular, note that integrating the differential equation (25) with $i = I$ between a generic point $z \in (z_{I-1}, z_I)$ and $z_I$ and imposing the condition $p(z_I) = 0$ (cf. the conditions (15)-(16)) we obtain

$$p(z) = \frac{c}{\mu_I}\,(z_I - z), \quad z \in (z_{I-1}, z_I). \tag{26}$$

In summary, $p(z)$ is positive, continuous, and monotonically decreasing on $(z_{i-1}, z_i)$ for all $i = 1, \ldots, I$. This along with the relations (21c) allow us to conclude that also $n_i(z)$ is positive, continuous, and monotonically decreasing on $(z_{i-1}, z_i)$ for all $i = 1, \ldots, I$.

*Step 2.* For $c \in \mathbb{R}_+^*$ given, under assumptions (6)-(7), introducing the additional notation

$$\alpha(z) := \sum_{i=1}^{I} \alpha_i\, \mathbb{1}_{(z_{i-1}, z_i)}(z), \tag{27}$$

and using the definition (23) of $\mu(z)$ along with the fact that, under conditions (15)-(16), both $p(z) = 0$ and $p'(z) = 0$ for all $z > z_I$, we further rewrite the set of differential equations (22) in a more compact form as

$$-c\,p'(z) - \big(\mu(z)\,p(z)\,p'(z)\big)' = \alpha(z)\,G(p)\,p(z), \quad z \in \mathbb{R}. \tag{28}$$

Multiplying by $p$ both sides of the differential equation (28) and rearranging terms we find

$$\mu\,p\,(p')^2 = \alpha\,G(p)\,p^2 + \big(p\,\mu\,p\,p'\big)' + c\,p\,p', \quad z \in \mathbb{R}. \tag{29}$$

Since the function $p(z)$ is continuous on $(z_{j-1}, z_j)$ for all $j = 1, \ldots, I-1$, there exists $z_j^* \in (z_{j-1}, z_j)$ such that $p'(z_j^*) > -\infty$ for any $j = 1, \ldots, I-1$. Hence, integrating both sides of (29) between $z_j^*$ and $z_{j+1}$ and estimating the right-hand side from above, by using the fact that the functions $p$, $\mu$, $\alpha$, and $G(p)$ are non-negative and bounded on $(z_{j-1}, z_{j+1})$ while the function $p'$ is non-positive on $(z_{j-1}, z_{j+1})$, we obtain

$$\int_{z_j^*}^{z_{j+1}} \mu\, p\, \left(p'\right)^2 \mathrm{d}z < \infty, \qquad j = 1, \ldots, I-1.$$

The above estimates ensure that $p' \in L^2_{loc}\left((z_j^*, z_{j+1})\right)$ for $j = 1, \ldots, I-1$. This along with the fact that $p \in L^\infty(\mathbb{R})$ allow us to conclude that $p$ is also continuous in each $z_i$ for $i = 1, \ldots, I-1$, i.e.

$$p(z_i^+) = p(z_i^-) = p(z_i), \quad i = 1, \ldots, I-1. \tag{30}$$

*Step 3.* For $c \in \mathbb{R}_+^*$ given, integrating the differential equation (28) between a generic point $z \in (z_{j-1}, z_{j+1})$ and $z_{j+1}$ for $j = 1, \ldots, I-1$, and using the fact that $p(z)$ is continuous and $\mu(z_{j+1}^-) = \mu_{j+1}$ (cf. the definition (23) of $\mu(z)$), yields

$$\begin{aligned} c\, p(z) + \mu(z)\, p(z)\, p'(z) = &\int_z^{z_{j+1}} \alpha\, G(p)\, p\, \mathrm{d}\zeta \\ &+ c\, p(z_{j+1}) + \mu_{j+1}\, p(z_{j+1})\, p'(z_{j+1}^-). \end{aligned} \tag{31}$$

Letting $z \to z_j^-$ in (31) and using the fact that $p(z)$ is continuous and $\mu(z_j^-) = \mu_j$ (cf. the definition (23) of $\mu(z)$) gives

$$\begin{aligned} c\, p(z_j) + \mu_j\, p(z_j)\, p'(z_j^-) = &\int_{z_j}^{z_{j+1}} \alpha\, G(p)\, p\, \mathrm{d}z \\ &+ c\, p(z_{j+1}) + \mu_{j+1}\, p(z_{j+1})\, p'(z_{j+1}^-). \end{aligned} \tag{32}$$

Similarly, letting $z \to z_j^+$ in (31) and using the fact that $p(z)$ is continuous and $\mu(z_j^+) = \mu_{j+1}$ (cf. the definition (23) of $\mu(z)$) gives

$$\begin{aligned} c\, p(z_j) + \mu_{j+1}\, p(z_j)\, p'(z_j^+) = &\int_{z_j}^{z_{j+1}} \alpha\, G(p)\, p\, \mathrm{d}z \\ &+ c\, p(z_{j+1}) + \mu_{j+1}\, p(z_{j+1})\, p'(z_{j+1}^-). \end{aligned} \tag{33}$$

Combining (32) and (33) we obtain

$$\mu_{i+1}\, p'(z_i^+) = \mu_i\, p'(z_i^-), \quad i = 1, \ldots, I-1. \tag{34}$$

*Step 4.* For $c \in \mathbb{R}_+^*$ given, combining (25) with $i = I$ (i.e. the fact that $c = -\mu_I\, p'(z_{I-1}^+)$) and the condition (34) with $i = I-1$ (i.e. the fact that $\mu_{I-1}\, p'(z_{I-1}^-) = \mu_I\, p'(z_{I-1}^+)$) we obtain

$$-\mu_{I-1}\, p'(z_{I-1}^-) = -\mu_I\, p'(z_{I-1}^+) = c. \tag{35}$$

Moreover, integrating the differential equation (24) between $z_i$ and $z_{i+1}$ for $i = 1, \dots, I-2$ and using the fact that $p(z)$ is continuous and $\mu(z_{i+1}^-) = \mu(z_i^+) = \mu_{i+1}$ (cf. the definition (23) of $\mu(z)$) yields

$$\begin{aligned} &-c\,(p(z_{i+1}) - p(z_i)) - \mu_{i+1}\left(p(z_{i+1})\, p'(z_{i+1}^-) - p(z_i)\, p'(z_i^+)\right) = 0, \\ &\quad i = 1, \dots, I-2, \end{aligned}$$

from which, rearranging terms, we find

$$p(z_i)\,\left(c + \mu_{i+1} p'(z_i^+)\right) - p(z_{i+1})\,\left(c + \mu_{i+1}\, p'(z_{i+1}^-)\right) = 0, \quad i = 1, \dots, I-2. \tag{36}$$

Choosing $i = I-2$ in (36) gives

$$p(z_{I-2})\,\left(c + \mu_{I-1} p'(z_{I-2}^+)\right) - p(z_{I-1})\,\left(c + \mu_{I-1}\, p'(z_{I-1}^-)\right) = 0$$

and then substituting (35) into the above equation yields

$$p(z_{I-2})\,\left(c + \mu_{I-1}\, p'(z_{I-2}^+)\right) = 0 \quad \Longrightarrow \quad -\mu_{I-1}\, p'(z_{I-2}^+) = c,$$

from which, using condition (34) with $i = I-2$ (i.e. the fact that $\mu_{I-1}\, p'(z_{I-2}^+) = \mu_{I-2}\, p'(z_{I-2}^-)$), we obtain

$$-\mu_{I-2}\, p'(z_{I-2}^-) = -\mu_{I-1}\, p'(z_{I-2}^+) = c. \tag{37}$$

Proceeding in a similar way for $i = I-3$ and so on and so forth for $i = I-4, \dots, 1$ we also find

$$-\mu_i\, p'(z_i^-) = -\mu_{i+1}\, p'(z_i^+) = c, \quad i = 1, \dots, I-3. \tag{38}$$

Taken together, the results given by (35), (37), and (38) allow us to conclude that

$$-\mu_{i+1}\, p'(z_i^+) = -\mu_i\, p'(z_i^-) = c, \quad i = 1, \dots, I-1, \tag{39}$$

and thus conditions (20) hold.

*Step 5.* For $c \in \mathbb{R}_+^*$ given, the expression (26) for $p$ on $(z_{I-1}, z_I)$ along with the constitutive relation (21c) with $i = I$ yield

$$n_I(z) = \frac{c}{\omega_I\, \mu_I}(z_I - z), \quad z \in (z_{I-1}, z_I). \tag{40}$$

Substituting (40) into (18) with $i = I$, and taking the positive root so as to ensure that the condition $z_I - z_{I-1} > 0$ (i.e. $z_I > z_{I-1}$) holds, gives

$$\frac{c}{\omega_I \mu_I} \int_{z_{I-1}}^{z_I} (z_I - z)\, \mathrm{d}z = M_I \quad \Longrightarrow \quad z_I = z_{I-1} + \sqrt{\frac{2\,\omega_I \mu_I}{c} M_I}. \tag{41}$$

Finally, substituting the expression (41) for $z_I$ into (26) and evaluating the resulting expression for $p(z)$ in $z = z_{I-1}$ yields

$$p(z_{I-1}^+) = \sqrt{\frac{2\,c\,\omega_I}{\mu_I} M_I},$$

from which, exploiting the fact that $p(z)$ is continuous, we obtain

$$p(z_{I-1}) = \sqrt{\frac{2\,c\,\omega_I}{\mu_I} M_I}. \tag{42}$$

*Step 6.* For $c \in \mathbb{R}_+^*$ given, integrating the differential equation (25) between a generic point $z \in (z_{i-1}, z_i)$ and $z_i$ for $i = 2, \ldots, I-1$, we find

$$p(z) = p(z_i) + \frac{c}{\mu_i}(z_i - z), \quad z \in (z_{i-1}, z_i), \quad i = 2, \ldots, I-1. \tag{43}$$

Choosing $i = I - 1$ in (43) gives

$$p(z) = p(z_{I-1}) + \frac{c}{\mu_{I-1}}(z_{I-1} - z), \quad z \in (z_{I-2}, z_{I-1}). \tag{44}$$

Integrating (44) between $z_{I-2}$ and $z_{I-1}$, using the fact that $p(z) = \omega_{I-1} n_{I-1}(z)$ for $z \in (z_{I-2}, z_{I-1})$ (cf. the constitutive relation (21c) with $i = I - 1$), and imposing the condition (18) with $i = I - 1$ gives

$$\int_{z_{I-2}}^{z_{I-1}} \left( p(z_{I-1}) + \frac{c}{\mu_{I-1}}(z_{I-1} - z) \right) \mathrm{d}z = \omega_{I-1} M_{I-1},$$

from which, computing the integral, solving the resulting quadratic equation for $z_{I-1} - z_{I-2}$, and taking the positive root so as to ensure that the condition $z_{I-1} - z_{I-2} > 0$ (i.e. $z_{I-1} > z_{I-2}$) holds, we find

$$z_{I-1} - z_{I-2} = \sqrt{\left(\frac{\mu_{I-1}}{c}\right)^2 (p(z_{I-1}))^2 + 2\,\frac{\mu_{I-1}}{c}\,\omega_{I-1} M_{I-1}} - \frac{\mu_{I-1}}{c}\, p(z_{I-1}). \tag{45}$$

Moreover, evaluating (44) in $z_{I-2}$ and substituting (45) into the resulting equation yields

$$p(z_{I-2}) = \frac{c}{\mu_{I-1}}\sqrt{\left(\frac{\mu_{I-1}}{c}\right)^2 (p(z_{I-1}))^2 + 2\,\frac{\mu_{I-1}}{c}\,\omega_{I-1}\,M_{I-1}} = \sqrt{(p(z_{I-1}))^2 + \frac{2\,c\,\omega_{I-1}}{\mu_{I-1}}\,M_{I-1}}.$$

Finally, substituting the expression (42) for $p(z_{I-1})$ into the above equation we obtain

$$p(z_{I-2}) = \sqrt{\frac{2\,c\,\omega_I}{\mu_I}\,M_I + \frac{2\,c\,\omega_{I-1}}{\mu_{I-1}}\,M_{I-1}}. \tag{46}$$

Proceeding in a similar way for $i = I-2$ and so on and so forth for $i = I-3, \ldots, 2$ we also find

$$p(z_i) = \sqrt{2\,c\sum_{j=i+1}^{I}\frac{\omega_j}{\mu_j}\,M_j}, \quad i = 2, \ldots, I-2. \tag{47}$$

Taken together, the results given by (42), (46), and (47) allow us to conclude that

$$p(z_i) = \sqrt{2\,c\sum_{j=i+1}^{I}\frac{\omega_j}{\mu_j}\,M_j}, \quad i = 1, \ldots, I-1. \tag{48}$$

Choosing $i = 1$ in (48) and recalling that $z_1 := 0$ (cf. conditions (16)), we obtain

$$p(0) = \sqrt{2\,c\sum_{j=2}^{I}\frac{\omega_j}{\mu_j}\,M_j}, \tag{49}$$

which implies that condition (19) holds.

*Step 7.* Evaluating (43) in $z_{i-1}$ and solving for $z_i$ yields

$$z_i = z_{i-1} + \frac{\mu_i}{c}\left(p(z_{i-1}) - p(z_i)\right), \quad i = 2, \ldots, I-1. \tag{50}$$

Choosing $i = 2$ in (50) and recalling that $z_1 := 0$ (cf. conditions (16)) gives

$$z_2 = \frac{\mu_2}{c}\left(p(0) - p(z_2)\right)$$

and then, substituting into the above equation the expression (49) for $p(0)$ and the expression for $p(z_2)$ obtained by choosing $i = 2$ in (48), we find

$$z_2 = \frac{\mu_2}{c}\left(\sqrt{2\,c\sum_{j=2}^{I}\frac{\omega_j}{\mu_j}M_j} - \sqrt{2\,c\sum_{j=3}^{I}\frac{\omega_j}{\mu_j}M_j}\right). \tag{51}$$

Moreover, proceeding in a similar way for $i = 3$ and so on and so forth for $i = 4, \ldots, I-1$ we also obtain

$$z_i = z_{i-1} + \frac{\mu_i}{c}\left(\sqrt{2\,c\sum_{j=i}^{I}\frac{\omega_j}{\mu_j}M_j} - \sqrt{2\,c\sum_{j=i+1}^{I}\frac{\omega_j}{\mu_j}M_j}\right), \quad i = 3, \ldots, I-1. \tag{52}$$

Finally, from (41) we find

$$z_I = z_{I-1} + \sqrt{\frac{2\,\omega_I\,\mu_I}{c}}M_I, \tag{53}$$

with $z_{I-1}$ obtained from (52) by choosing $i = I-1$.
*Step 8.* Complementing the differential equation (22a) with the condition (17) and the condition (39) for $i = 1$, we obtain the following problem

$$\begin{cases} -c\,p'(z) - \mu_1\,\left(p(z)\,p'(z)\right)' = \alpha_1\,G(p)\,p(z), \quad z \in (z_0, z_1) \equiv (-\infty, 0), & \text{(54a)} \\ p(-\infty) = \overline{p}, \quad p'(0^-) = -\dfrac{c}{\mu_1}, & \text{(54b)} \end{cases}$$

from which, proceeding as similarly done in Lorenzi et al. (2017), it is possible to prove (see Appendix B:) that $c \mapsto p(0^-)$ is monotonically decreasing. On the other hand, the expression (49) for $p(0)$ implies that $c \mapsto p(0^+)$ is monotonically increasing. These facts along with the condition $p(0^-) = p(0^+)$, which follows from the continuity of $p(z)$, allow us to conclude that there exists a unique pair $(c, p)$ that satisfies the travelling wave problem. □

***Remark 3*** The fact that, as established by Theorem 1, the cell pressure $p(z)$ defined via the constitutive relation (14c) is continuous throughout the support of the travelling wave gives the following interface conditions for the cell densities

$$n_{i+1}(z_i^+) = \frac{\omega_i}{\omega_{i+1}}n_i(z_i^-), \quad i = 1, \ldots, I-1. \tag{55}$$

These conditions imply that if $\omega_i < \omega_{i+1}$ then $n_{i+1}(z_i^+) < n_i(z_i^-)$, whereas if $\omega_i \geq \omega_{i+1}$ then $n_{i+1}(z_i^+) \geq n_i(z_i^-)$.

***Remark 4*** The conditions (20) imply that

$$-p'(z_i^+) = -\frac{\mu_i}{\mu_{i+1}}\, p'(z_i^-), \quad i = 1, \dots, I-1$$

and, therefore, since $p'(z_i^+) < 0$ and $p'(z_i^-) < 0$ for all $i = 1, \dots, I-1$, we have

$$|p'(z_i^+)| = \frac{\mu_i}{\mu_{i+1}}\, |p'(z_i^-)|, \quad i = 1, \dots, I-1.$$

Under assumptions (13), which imply that $\dfrac{\mu_i}{\mu_{i+1}} < 1$ for all $i = 1, \dots, I-1$, the above relations allow us to conclude that

$$|p'(z_i^+)| < |p'(z_i^-)|, \quad i = 1, \dots, I-1.$$

## 4 Numerical Simulations

In this section, we present results of numerical simulations of the individual-based model introduced in Sect. 2 and the corresponding continuum model defined by the PDE system (3), and compare them to the results of travelling wave analysis obtained in Sect. 3. In particular, we investigate the cases where there are either three or four different cellular phenotypes (i.e. $I = 3$ or $I = 4$).

### 4.1 Set-Up of Numerical Simulations

We carry out numerical simulations over the spatial domain $[0, L]$, with $L = 150$. In order to ensure that assumptions (6) are satisfied, we use the following definition of the function $G(p)$, both when $I = 3$ and when $I = 4$,

$$G(p) := \arctan\left(\frac{1}{10}\left(1 - \frac{p}{\overline{p}}\right)\right). \tag{56}$$

Furthermore, we choose values of the parameters $\alpha_i$ and $\mu_i$ satisfying assumptions (7) and (13). Specifically, in the case where $I = 3$ we use the following baseline parameter values

$$\begin{aligned} &\alpha_1 = 10, \quad \alpha_2 = \alpha_3 = 0, \\ &\mu_1 = 10^{-4}, \quad \mu_2 = 2 \times 10^{-4}, \quad \mu_3 = 3 \times 10^{-4}, \\ &\omega_1 = 1, \quad \omega_2 = 2, \quad \omega_3 = 3, \end{aligned} \tag{57}$$

and in the case where $I = 4$ we complement the parameter choice given by (57) with

$$\alpha_4 = 0, \quad \mu_4 = 4 \times 10^{-4}, \quad \omega_4 = 4. \tag{58}$$

In both cases, we explore also deviations of the weights $\omega_i$ from these values in the numerical simulations.

Moreover, in line with the travelling wave analysis carried out in Sect. 3, we investigate propagation of segregation properties by letting cells with different phenotypes be spatially segregated at $t = 0$. Specifically, both for the individual-based model and for the continuum model, when $I = 3$ we define the initial cell densities as

$$\begin{aligned} n_1(0,x) &= A_1 \exp\left[-B\,x^2\right] \mathbb{1}_{[0,10)}(x), \\ n_2(0,x) &= A_2 \exp\left[-B\,(x-10)^2\right] \mathbb{1}_{[10,20)}(x), \\ n_3(0,x) &= A_3 \exp\left[-B\,(x-20)^2\right] \mathbb{1}_{[20,L)}(x), \end{aligned} \tag{59}$$

while when $I = 4$ we use the following definitions for the initial cell densities

$$\begin{aligned} n_1(0,x) &= A_1 \exp\left[-B\,x^2\right] \mathbb{1}_{[0,10)}(x), \\ n_2(0,x) &= A_2 \exp\left[-B\,(x-10)^2\right] \mathbb{1}_{[10,20)}(x), \\ n_3(0,x) &= A_3 \exp\left[-B\,(x-20)^2\right] \mathbb{1}_{[20,30)}(x), \\ n_4(0,x) &= A_4 \exp\left[-B\,(x-30)^2\right] \mathbb{1}_{[30,L)}(x). \end{aligned} \tag{60}$$

In (59) and (60), the function $\mathbb{1}_{(\cdot)}(x)$ is the indicator function of the set $(\cdot)$, $B = 6 \times 10^{-2}$, and the parameters $A_i$ are positive real numbers. We choose the homeostatic pressure to be

$$\overline{p} = 4 \times 10^4$$

and, given this value of $\overline{p}$ and the values selected for the weights $\omega_i$ in (3c) and (5), we then choose the values of the parameters $A_i$ in (59) and (60) such that the initial cell pressure satisfies $p(0,x) \leq \overline{p}$ for all $x \in [0, L]$ and is consistent between numerical simulations.

Finally, to carry out numerical simulations of the individual-based model, we choose the space-step $\Delta_x = 0.1$, the time-step $\tau = 1 \times 10^{-4}$, and we define

$$\gamma_i := \frac{2\tau\overline{p}}{\Delta_x^2}\mu_i \tag{61}$$

so as to ensure that conditions (11) underlying the formal derivation of the continuum model are met. Note that, since we choose values of the parameters $\mu_i$ satisfying assumptions (13), defining the values of the parameter $\gamma_i$ according to (61) ensures that assumptions (9) are satisfied as well.

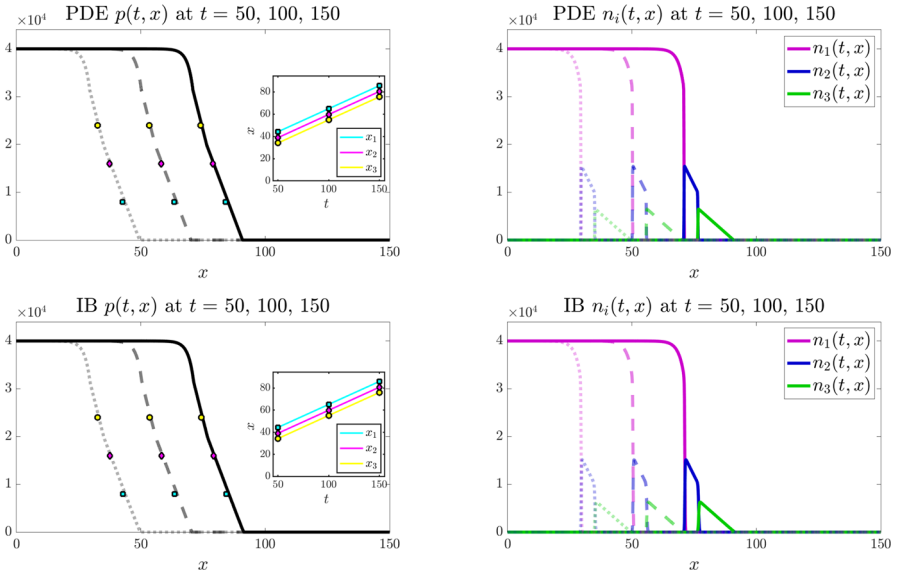


**Fig. 3** Main results under the baseline parameter setting for $I = 3$. Comparing numerical solutions of the continuum model (top panels) with the averaged results of 10 simulations of the individual-based model (bottom panels), when $I = 3$ and the values of the parameters $\alpha_i$, $\mu_i$, and $\omega_i$ are set according to (57). Plots display the cell pressure $p(t, x)$ (left panels) and the cell densities $n_i(t, x)$ (right panels) at three successive time instants – i.e. $t = 50$ (dotted lines), $t = 100$ (dashed lines), and $t = 150$ (solid lines). The insets of the left panels display the plots of $x_1(t)$ (cyan), $x_2(t)$ (magenta), and $x_3(t)$ (yellow) defined via (62). The coloured markers in the plot of $p(t, x)$ highlight the values of $p(t, x_1(t))$ (cyan), $p(t, x_2(t))$ (magenta), and $p(t, x_3(t))$ (yellow) at $t = 50$, $t = 100$, and $t = 150$. The numerically estimated wave speeds are $c_{\text{PDEn}} = 0.42$ and $c_{\text{IBn}} = 0.42$, and the analytically predicted wave speed is $c_{\text{a}} = 0.42$ (Color figure online)

### 4.2 Computational Implementation of the Individual-Based Model and Numerical Scheme for the Continuum Model

All simulations are performed in MATLAB. For the individual-based model, at each time-step, every individual cell can undergo: (i) movement, according to the probabilities defined in Sect. 2.1.2; (ii) division and death, according to the probabilities defined in Sect. 2.1.1. For each of these processes a random number is drawn from the standard uniform distribution on the interval (0, 1) using the built-in MATLAB function RAND. If this random number is smaller than the probability of the event occurring then the process is successful. To impose zero-flux boundary conditions, any attempted move outside the spatial domain is aborted. To numerically solve the PDE system (3) subject to zero-flux boundary conditions we use a finite volume scheme modified from our previous works (Bubba et al. 2020; Lorenzi et al. 2023). The full details of this scheme are provided in Appendix C.

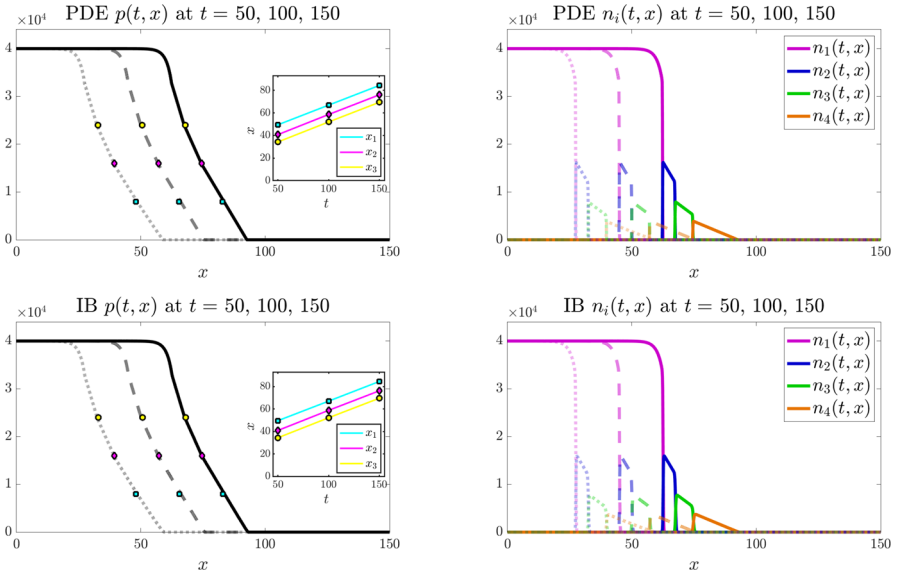


**Fig. 4** Main results under the baseline parameter setting for $I = 4$. Comparing numerical solutions of the continuum model (top panels) with the averaged results of 10 simulations of the individual-based model (bottom panels), when $I = 4$ and the values of the parameters $\alpha_i$, $\mu_i$, and $\omega_i$ are set according to (57)-(58). Plots display the cell pressure $p(t, x)$ (left panels) and the cell densities $n_i(t, x)$ (right panels) at three successive time instants – i.e. $t = 50$ (dotted lines), $t = 100$ (dashed lines), and $t = 150$ (solid lines). The insets of the left panels display the plots of $x_1(t)$ (cyan), $x_2(t)$ (magenta), and $x_3(t)$ (yellow) defined via (62). The coloured markers in the plot of $p(t, x)$ highlight the values of $p(t, x_1(t))$ (cyan), $p(t, x_2(t))$ (magenta), and $p(t, x_3(t))$ (yellow) at $t = 50$, $t = 100$, and $t = 150$. The numerically estimated wave speeds are $c_{\text{PDEn}} = 0.35$ and $c_{\text{IBn}} = 0.35$, and the analytically predicted wave speed is $c_{\text{a}} = 0.35$ (Color figure online)

### 4.3 Main Results of Numerical Simulations

Figures 3 and 4 display the results of numerical simulations of the individual-based model (bottom panels) and the corresponding continuum model defined by the PDE system (3) (top panels) for $I = 3$ and $I = 4$, respectively, under the baseline parameter settings (57) and (57)-(58), respectively. Moreover, Figs. 6 and 7 display the results of numerical simulations of the continuum model for $I = 3$ and $I = 4$, respectively, when the values of the parameters $\alpha_i$ and $\mu_i$ are set according to (57) and (57)-(58), respectively, while different combinations of the weights $\omega_i$ are considered.

#### 4.3.1 Agreement between the individual-based and the continuum models

The results summarised by the plots in Figs. 3 and 4 demonstrate that overall there is excellent agreement between numerical simulations of the individual-based model and numerical solutions of the corresponding continuum model, as expected since conditions (11) are met. Moreover, although, for the sake of clarity, only the results of numerical simulations of the continuum model are displayed in Figs. 6 and 7, we verified that also for the parameter settings corresponding to these figures the results of numerical simulations of the individual-based model agree with those of the continuum

model (results not shown). We stress that possible quantitative discrepancies between the two models emerge in the proximity of the interfaces between regions occupied by cells with different phenotypes (see Fig. 5). This is because in these areas there is a stronger interplay between demographic stochasticity and sharp transitions in cell densities, which causes a reduction in the quality of the continuum approximations that are employed in the formal derivation of the PDE model from the underlying individual-based model. We also remark that the quantitative agreement between the two models may deteriorate in scenarios where a limited number of cells is considered, since in these scenarios the impact of demographic stochasticity is amplified.

#### 4.3.2 Propagation of travelling fronts

The numerical results in Figs. 3-4 and Figs. 6-7 show the propagation of travelling fronts wherein, as expected from Theorem 1, the cell densities $n_i$ have disjoint supports, which means that cells with different phenotypes occupy distinct regions across the front (cf. right panels). In particular, as captured by assumptions (6)-(7) and (13), cells with phenotypes labelled by larger values of the index $i$, which display a higher migratory ability, occupy regions closer to the edge of the front, while fast-dividing cells with the lowest mobility (i.e. cells with phenotype $i = 1$) make up the bulk of the population in the rear of the front. Moreover, also in agreement with Theorem 1, the cell pressure $p$ is continuous throughout the wave, whereas its first spatial derivative exhibits jump discontinuities at the interfaces between the regions occupied by cells with different phenotypes (cf. left panels). Specifically, after an initial transient during which the travelling front is formed, the numerical values of the first spatial derivative of the cell pressure $p$ are such that the interface conditions (20) are satisfied.

In order to confirm the propagation of travelling waves, we also track the dynamics of the points $x_1(t)$, $x_2(t)$, and $x_3(t)$ such that

$$p(t, x_1(t)) = 0.2\,\overline{p}, \quad p(t, x_2(t)) = 0.4\,\overline{p}, \quad p(t, x_3(t)) = 0.6\,\overline{p}, \tag{62}$$

and verify that, after an initial transient during which the travelling wave is formed, the functions $x_1(t)$, $x_2(t)$, and $x_3(t)$ behave like straight lines with approximately the same constant slope (cf. insets of the left panels). We numerically estimate the wave speeds for the individual-based model and for the continuum model, denoted $c_{\mathrm{IBn}}$ and $c_{\mathrm{PDEn}}$, respectively, by measuring the slope of $x_1(t)$ after the transient. We then compare the numerically estimated wave speeds with the analytically predicted one, denoted $c_{\mathrm{a}}$, which is obtained through (19), i.e. substituting into the following formula

$$c_{\mathrm{a}} = \frac{(p(0))^2}{2\sum_{j=2}^{I} \frac{\omega_j}{\mu_j} M_j}$$

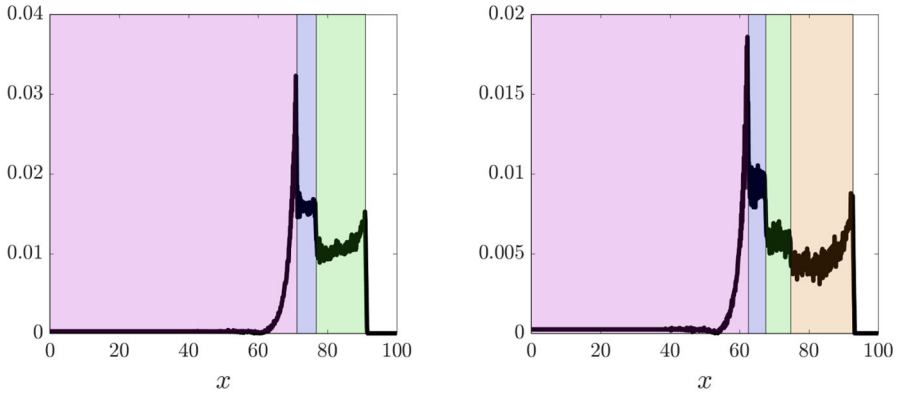


**Fig. 5** Quantitative comparison between the individual-based and the continuum models. Plot of the quantity $\frac{|p_{PDE}(t,x) - p_{IB}(t,x)|}{\overline{p}}$ at $t = 150$, where $p_{PDE}$ is the cell pressure computed from numerical solutions of the continuum model displayed in Fig. 3 (left panel) and Fig. 4 (right panel), while $p_{IB}$ is the cell pressure computed from the averaged results of 10 simulations of the individual-based model displayed in the same figures. The supports of the cell densities $n_i$ for $i = 1, \ldots, I$ with $I = 3$ (left panel) or $I = 4$ (right panel) are highlighted in the same colours as those of the curves of the cell densities displayed in Figs. 3 and 4 (Color figure online)

the values of $M_2, \ldots, M_I$ and the value of $p(0)$ estimated from numerical solutions of the continuum model. In particular, we approximate $M_j$ with the numerical value of the integral of the cell density $n_j$ over the spatial domain for $j = 2, \ldots, I$, which remains constant over time, while $p(0)$ is approximated as the numerical value of the cell pressure $p$ at the right endpoint of the support of the cell density $n_1$ at a time $t$ large enough that the travelling wave is established. We find that there is good agreement between the values of $c_{\mathrm{IBn}}$, $c_{\mathrm{PDEn}}$, and $c_{\mathrm{a}}$ (cf. the values provided in the captions of Figs. 3, 4, 6 and 7). We also verified that, when the travelling wave is fully formed, the positions of the right endpoints of the supports of the cell densities are consistent with those predicted by the travelling wave analysis (i.e. those given by (51)-(53)). In more detail, denoting by $X_i$ the position of the right endpoint of the support of the cell density $n_i$ at time $t$ large enough that the travelling wave is established, recalling that $z_1 := 0$ (cf. conditions (16)), we verified that the values of $Z_i := X_i - X_1$ for $i = 2, \ldots, I$ are consistent with the analytically predicted values, denoted by $Z_{ai}$, which are obtained by substituting into (51)-(53) the analytically predicted wave speed, $c_{\mathrm{a}}$, and the numerically estimated values of $M_j$ for $j = 2, \ldots, I$.

***Impact of the parameters $\omega_i$ on the shape of travelling fronts***

The numerical results in Figs. 6 and 7 show that, in agreement with the analytical results of Theorem 1, the choice of the values of the parameters $\omega_i$ impacts on the shape of the travelling fronts that emerge. In more detail, these numerical results indicate that, once the travelling front is established, the cell densities $n_i$ are such that the interface conditions (55) are satisfied throughout the front. Hence, at the interface between the region occupied by cells with phenotype labelled by the index $i + 1$ and the region occupied by cells with phenotype labelled by the index $i$ (i.e. at the interface between the supports of $n_{i+1}$ and $n_i$), for $i = 1, \ldots, I - 1$: if $\omega_{i+1} > \omega_i$ then the value of the

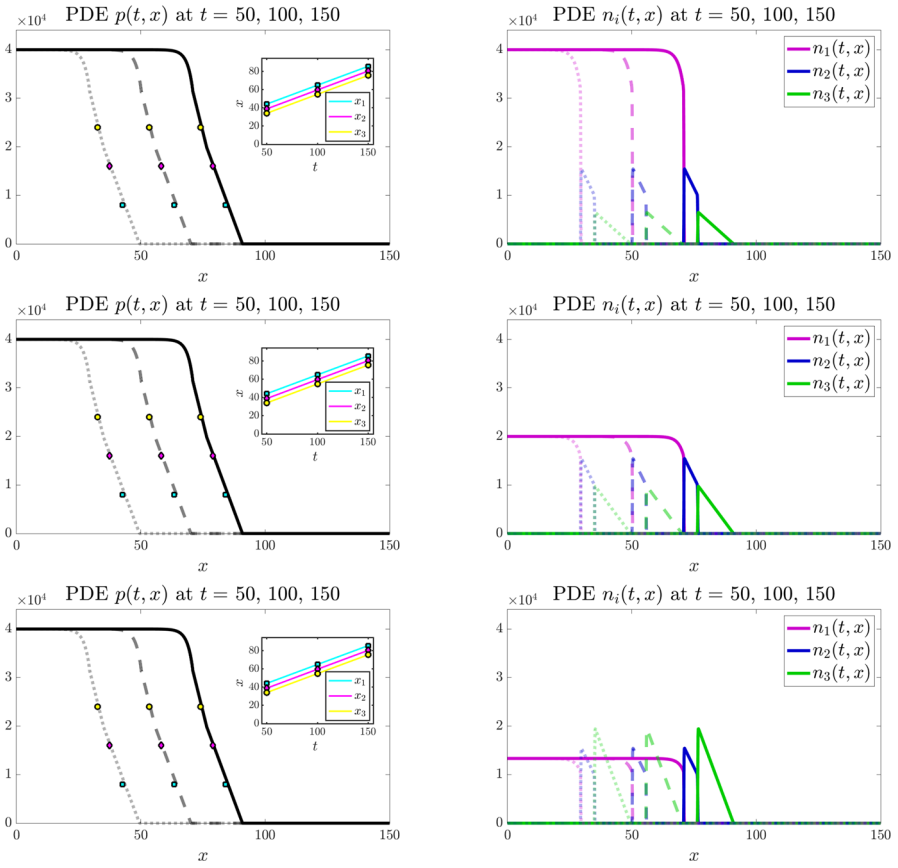


**Fig. 6** Main results under different values of the parameters $\omega_i$ for $I = 3$. Numerical solutions of the continuum model when $I = 3$ and the values of the parameters $\alpha_i$ and $\mu_i$ are set according to (57), while the values of the parameters $\omega_i$ are: $\omega_1 = 1,\ \omega_2 = 2,\ \omega_3 = 3$ (top panels); $\omega_1 = 2,\ \omega_2 = 2,\ \omega_3 = 2$ (central panels); and $\omega_1 = 3,\ \omega_2 = 2,\ \omega_3 = 1$ (bottom panels). Plots display the cell pressure $p(t,x)$ (left panels) and the cell densities $n_i(t,x)$ (right panels) at three successive time instants – i.e. $t = 50$ (dotted lines), $t = 100$ (dashed lines), and $t = 150$ (solid lines). The insets of the left panels display the plots of $x_1(t)$ (cyan), $x_2(t)$ (magenta), and $x_3(t)$ (yellow) defined via (62). The coloured markers in the plot of $p(t,x)$ highlight the values of $p(t,x_1(t))$ (cyan), $p(t,x_2(t))$ (magenta), and $p(t,x_3(t))$ (yellow) at $t = 50$, $t = 100$, and $t = 150$. For all cases we report on in this figure, the numerically estimated wave speed is $c_{\mathrm{PDEn}} = 0.42$, while the analytically predicted wave speed is $c_{\mathrm{a}} = 0.42$ (Color figure online)

cell density at the right of the interface is smaller than the one at the left (cf. top, right panels); if $\omega_{i+1} = \omega_i$ then the value of the cell density at the right of the interface is the same as the one at the left (cf. central, right panels); if $\omega_{i+1} < \omega_i$ then the value of the cell density at the right of the interface is larger than the one at the left (cf. bottom, right panels).

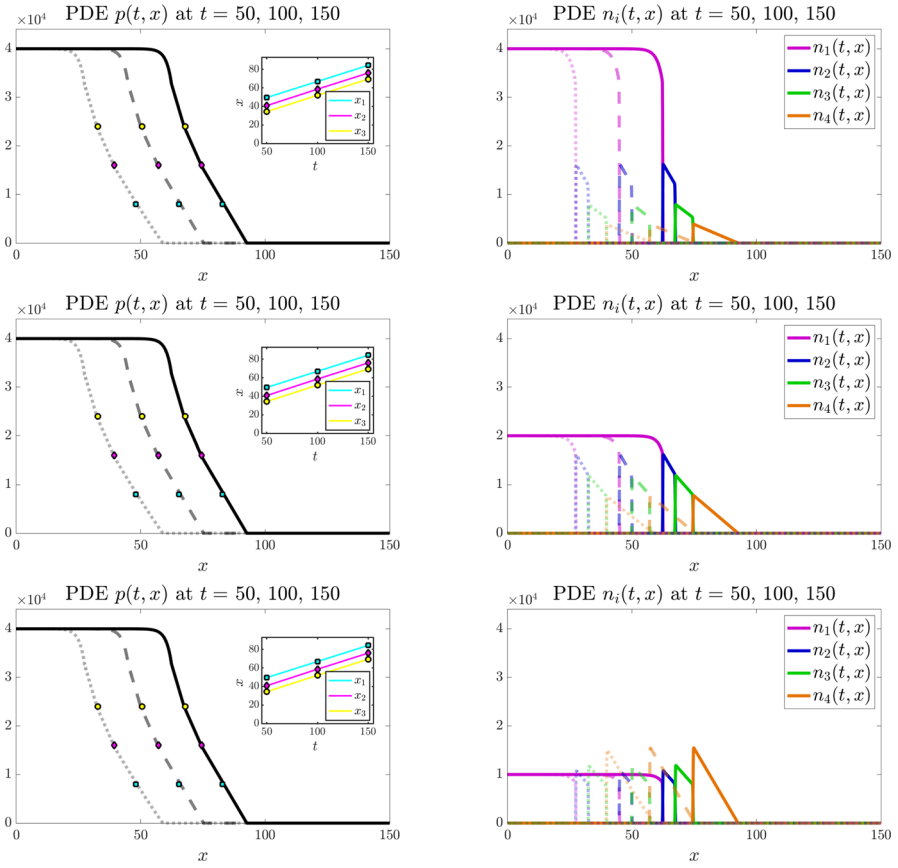


**Fig. 7** Main results under different values of the parameters $\omega_i$ for $I = 4$. Numerical solutions of the continuum model when $I = 4$ and the values of the parameters $\alpha_i$ and $\mu_i$ are set according to (57)-(58), while the values of the parameters $\omega_i$ are: $\omega_1 = 1,\ \omega_2 = 2,\ \omega_3 = 3,\ \omega_4 = 4$ (top panels); $\omega_1 = 2,\ \omega_2 = 2,\ \omega_3 = 2,\ \omega_4 = 2$ (central panels); and $\omega_1 = 4,\ \omega_2 = 3,\ \omega_3 = 2,\ \omega_4 = 1$ (bottom panels). Plots display the cell pressure $p(t,x)$ (left panels) and the cell densities $n_i(t,x)$ (right panels) at three successive time instants – i.e. $t = 50$ (dotted lines), $t = 100$ (dashed lines), and $t = 150$ (solid lines). The insets of the left panels display the plots of $x_1(t)$ (cyan), $x_2(t)$ (magenta), and $x_3(t)$ (yellow) defined via (62). The coloured markers in the plot of $p(t,x)$ highlight the values of $p(t,x_1(t))$ (cyan), $p(t,x_2(t))$ (magenta), and $p(t,x_3(t))$ (yellow) at $t = 50$, $t = 100$, and $t = 150$. For all cases we report on in this figure, the numerically estimated wave speed is $c_{\mathrm{PDEn}} = 0.35$, while the analytically predicted wave speed is $c_{\mathrm{a}} = 0.35$ (Color figure online)

## 5 Discussion and Research Perspectives

In this work, we have considered a PDE model for the growth of heterogeneous cell populations subdivided into multiple distinct discrete phenotypes. In this model, cells preferentially move towards regions where they are less compressed, and thus their movement occurs down the gradient of the cellular pressure. The cellular pressure is defined as a weighted sum of the densities (i.e. the volume fractions) of cells with different phenotypes. To translate into mathematical terms the idea that cells with dif-

ferent phenotypes have different morphological and mechanical properties, both the cell mobility and the weighted amount the cells contribute to the cellular pressure vary with their phenotype. We have formally derived this model as the continuum limit of an on-lattice individual-based model, where cells are represented as single agents undergoing a branching biased random walk corresponding to phenotype-dependent and pressure-regulated cell division, death, and movement. Then, we have studied travelling wave solutions whereby cells with different phenotypes are spatially separated by sharp boundaries across the invading front (cf. Theorem 1, Remark 2, and the schematic in Fig. 2). As discussed in Batlle and Wilkinson (2012), sharp borders between distinct cell types form at the interface of both adjacent tissues and regional domains within a tissue, and the occurrence of sharp spatial segregation between cells is observed both in *Drosophila* (Irvine 2001) and in vertebrate tissues (Fraser et al. 1990; Langenberg and Brand 2005; Zeltser and Larsen 2001). Finally, we have reported on numerical simulations of the two models, demonstrating excellent agreement between them and the travelling wave analysis, thus validating the formal limiting procedure employed to derive the individual-based model from the continuum model and confirming the analytical results obtained.

The results presented here indicate that inter-cellular variability in mobility can support the maintenance of spatial segregation across invading fronts, whereby cells with a higher mobility drive invasion by occupying regions closer to the front edge. These results have been obtained under the assumption that cells with phenotypes labelled by larger values of the index $i$ express a higher mobility – cf. assumptions (9) on the parameters $\gamma_i$ of the individual-based model and the corresponding assumptions (13) on the parameters $\mu_i$ of the continuum model. On the other hand, no specific assumptions have been made on the parameters $\omega_i$, which provide a measure of the weighted amount that cells with phenotype $i$ contribute towards the cellular pressure (cf. the constitutive relations (3c) and (5)) and the values of which can be related, for instance, to cell stiffness – i.e. if cells with phenotype labelled by the index $i$ are stiffer than cells with phenotype labelled by the index $j$ then $\omega_i > \omega_j$. Therefore, the results we have presented apply both to scenarios where less stiff cells are more invasive and to opposite scenarios – scenarios that, under assumptions (9) and (13), would correspond to assuming, respectively, $\omega_{i+1} < \omega_i$ and $\omega_{i+1} > \omega_i$ for $i = 1, \dots, I-1$. This is particularly relevant in the context of tumour growth. In fact, it has been observed, through both *in vitro* and *in vivo* experiments, that solid tumours can be made up of cells of varying stiffness (Lv et al. 2021; Swaminathan et al. 2011; Rianna et al. 2020; Baker et al. 2010; Han et al. 2020). In a large number of cancer cell lines, more migratory and invasive phenotypes align with the cells that are softest/most deformable (Lv et al. 2021; Han et al. 2020). For example, in ovarian cancer cell lines, cells with the highest migration capabilities can be up to ten times less stiff than cells with the lowest migration and invasion potential (Swaminathan et al. 2011). However, in some cell lines, such as breast cancer cell lines, stiffness has been observed to increase with tumourigenic invasive potential, especially in areas of stiff extracellular matrix (Baker et al. 2010; Mok et al. 2020).

We conclude with an outlook on possible research perspectives. In this work the focus has been placed on the *propagation* of segregation properties, and we have thus studied the existence of travelling wave solutions of the system of PDEs (3)

that exhibit segregation between different cell types. Accordingly, we have carried out numerical simulations under initial conditions corresponding to scenarios where cells of different types initially occupy distinct regions of the spatial domain. As the next step, it would be relevant to investigate the *emergence* of segregation properties by studying analytically the convergence of solutions of the PDE system (3) to such travelling wave solutions, and carrying out numerical simulations in cases where cells of different types are not separated at the initial time.

It would also be interesting to explore scenarios where assumptions (7) on the parameters $\alpha_i$ are relaxed (i.e. when also proliferation and death of cells with phenotypes labelled by values of the index $i > 1$ are incorporated into the model). In this regard, under assumptions (13) on the mobility coefficients $\mu_i$, taking into account proliferation-migration trade-offs induced by the inherent energetic cost attached to cellular activities, it would be natural to consider the variant (12) of the PDE system (3) subject to assumptions (6) along with the assumptions $\alpha_1 > \alpha_2 > \ldots > \alpha_{I-1} > \alpha_I = 0$. In this case, the calculations carried out in *Step 4* of the proof of Theorem 1 would break down, hinting that conditions (20), which express the fact that the interfaces between the components $n_i(z)$ and $n_{i+1}(z)$ of the solution travel at the same speed $c$ for all $i = 1, \ldots, I-1$, would not hold. Hence, under this scenario it is likely that the continuum model does not admit travelling wave solutions of the type of those of Theorem 1.

While, focusing our attention on spatial segregation across travelling fronts, here we have considered the case where assumptions (13) hold, this work could be extended by investigating the behaviour of solutions to the PDE system (3), and its more general variant (12), in cases where these assumptions do not hold and segregation properties may not be propagated (David et al. 2024; Lorenzi et al. 2017; Carrillo 2018a). It would also be interesting to study free boundary problems for these PDE systems, along the lines of those considered in Byrne and Chaplain (1997); Lorenzi et al. (2020); Bertsch et al. (2010) to model tissue development and tumour growth, and consider related transmission problems modelling cell invasion through thin membranes, in the vein of Chaplain (2019); Ciavolella et al. (2024); Ciavolella and Perthame (2021); Giverso et al. (2022).

Moreover, we could incorporate into the PDE system (3), or its generalised variant (12), the effects of phenotypic switching, possibly driven by the extracellular environment (Celora et al. 2021; Charras and Sahai 2014), and cell-cell adhesion, as similarly done for instance in Bubba et al. (2020); Carrillo (2018b); Crossley et al. (2024); Macfarlane et al. (2022) and Berendsen et al. (2017); Burger et al. (2020); Carrillo (2019, 2018a, b), respectively. Instead of resorting to phenomenological considerations to define the additional terms modelling these phenomena, we could extend the individual-based model considered here along with the formal procedure employed to derive the corresponding continuum model so as to encompass mechanisms of phenotypic switching and cell-cell adhesion. This would be biologically interesting in that it has been shown, on the one hand, that some cancer cell lines can adaptively alter their stiffness to become softer in order to overcome physical barriers or areas of high cellular pressure (Rianna et al. 2020; Han et al. 2020), and, on the other hand, that in general internal stiffness of aggressive tumours is more heterogeneous than in quiescent tumours (Mok et al. 2020). Adapting our current model to include dynamical

variation in cell stiffness through the weights modelling contribution to the cellular pressure could be of interest to study these dynamics further. On this note, recent work by Zills et al. (2023) considered a 3D individual-based model of a growing cell population where cells could become softer during cell division. The cell stiffness then related to adhesive and repulsive forces of the cells and their contribution to the local cellular pressure, which in turn regulated cell division rates. Numerical simulations of this model were able to replicate experimental data on cellular spheroids where cells towards the border were faster, larger, and softer than those at the centre of the spheroid. However, as the model considered is an individual-based model, the study in Zills et al. (2023) is based on numerical simulations only. Hence, it would be interesting to extend our modelling framework to include such additional aspects and then carry out travelling wave analysis, in order to complement numerical simulations with analytical results to facilitate a more comprehensive exploration of the model parameter space.

Another avenue for future research could be to investigate spatial segregation across travelling fronts in individual-based and continuum models for the growth of heterogeneous cell populations that encapsulate volume exclusion effects, which are not captured by the models considered here. For this, we expect modelling methods and analytical techniques similar to those employed in Baker et al. (2019); Fadai and Simpson (2020); Murphy et al. (2020, 2021); Tambyah et al. (2020) to be useful.

As a further extension of the present work, building on the modelling approach presented in David (2023); Lorenzi et al. (2023); Lorenzi and Painter (2022); Lorenzi et al. (2021), we could also let the cell phenotype vary along a spectrum, and thus be described by a continuous variable $y \in \mathbb{R}$ (Lorenzi et al. 2024). In this case, the evolution of the density (i.e. the volume fraction) of cells with phenotype $y$ at time $t \geq 0$, $n(t, x, y)$, would be governed by the following partial integro-differential equation model

$$\begin{cases} \partial_t n - \mu(y)\, \partial_x \left(n\, \partial_x p\right) = \alpha(y)\, G(p)\, n, \quad y \in \mathbb{R}, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (t, x) \in (0, \infty) \times \mathbb{R}. \\ p(t, x) := \displaystyle\int_{\mathbb{R}} \omega(y)\, n(t, x, y)\, \mathrm{d}y, \end{cases} \tag{63}$$

Note that, compared to the generalised form (12) of the PDE model (3), subject to assumptions (6), here the parameters $\mu_i$, $\alpha_i$, and $\omega_i$ have been replaced by the functions $\mu(y)$, $\alpha(y)$, and $\omega(y)$, respectively. We expect the model (63) to be derivable from an underlying individual-based model through a formal limiting procedure analogous to the one employed in Lorenzi et al. (2023); Macfarlane (2022).

The aforementioned extensions of the present work, which will bring new mathematical problems, will allow for further investigation into how phenotypic heterogeneity shapes invading fronts in growing cell populations.

## Appendix A: Formal Derivation of the Continuum Model from the Individual-Based Model

We detail the formal derivation of the PDE system (3) from the branching biased random walk underlying the individual-based model described in Sect. 2.

When cell dynamics are governed by the rules described in Sect. 2.1, using that $1 + \tau G_i(p_j^k)_+ - \tau G_i(p_j^k)_- = 1 + \tau G_i(p_j^k)$, the principle of mass balance gives the balance equation (10). Then, employing a formal procedure analogous to the one that we used in Chaplain et al. (2020); Macfarlane and Chaplain (2020); Macfarlane (2022), we define

$$n_i \equiv n_i(t,x) := n_{i,\,j}^k, \quad n_i(t+\tau,x) := n_{i,\,j}^{k+1}, \quad n_i(t,x\pm\Delta_x) := n_{i,\,j\pm1}^k$$

and

$$p \equiv p(t,x) := p_j^k, \quad p(t,x\pm\Delta_x) := p_{j\pm1}^k.$$

From the constitutive relation (5) we find

$$p(t,x) := \sum_{i=1}^{I} \omega_i\, n_i(t,x), \tag{A1}$$

while from the balance equation (10) we obtain

$$\begin{aligned}
&n_i(t+\tau,x)\\
&\quad = n_i\left\{(1+\tau G_i(p))\left[1-\frac{\gamma_i\,(p-p(t,x+\Delta_x))_+}{2\overline{p}}-\frac{\gamma_i\,(p-p(t,x-\Delta_x))_+}{2\overline{p}}\right]\right\}\\
&\qquad +n_i(t,x+\Delta_x)\left\{(1+\tau G_i(p(t,x+\Delta_x)))\left[\frac{\gamma_i\,(p(t,x+\Delta_x)-p)_+}{2\overline{p}}\right]\right\}\\
&\qquad +n_i(t,x-\Delta_x)\left\{(1+\tau G_i(p(t,x-\Delta_x)))\left[\frac{\gamma_i\,(p(t,x-\Delta_x)-p)_+}{2\overline{p}}\right]\right\}.
\end{aligned}$$

Splitting the growth terms gives

$$\begin{aligned}
n_i(t+\tau,x) = n_i&\left[1-\frac{\gamma_i\,(p-p(t,x+\Delta_x))_+}{2\overline{p}}-\frac{\gamma_i\,(p-p(t,x-\Delta_x))_+}{2\overline{p}}\right]\\
&+n_i(t,x+\Delta_x)\left[\frac{\gamma_i\,(p(t,x+\Delta_x)-p)_+}{2\overline{p}}\right]\\
&+n_i(t,x-\Delta_x)\left[\frac{\gamma_i\,(p(t,x-\Delta_x)-p)_+}{2\overline{p}}\right]\\
&+\tau G_i(p)n_i\left[1-\frac{\gamma_i\,(p-p(t,x+\Delta_x))_+}{2\overline{p}}-\frac{\gamma_i\,(p-p(t,x-\Delta_x))_+}{2\overline{p}}\right]\\
&+\tau G_i(p(t,x+\Delta_x))n_i(t,x+\Delta_x)\left[\frac{\gamma_i\,(p(t,x+\Delta_x)-p)_+}{2\overline{p}}\right]
\end{aligned}$$

$$+\tau G_i(p(t,x-\Delta_x))n_i(t,x-\Delta_x)\left[\frac{\gamma_i\,(p(t,x-\Delta_x)-p)_+}{2\overline{p}}\right].$$

Now, assuming $n_i$ to be sufficiently regular, using that

$$n_i(t,x\pm\Delta_x)=n_i\pm\Delta_x\frac{\partial n_i}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 n_i}{\partial x^2}+\mathcal{O}(\Delta_x^3),$$

we formally obtain

$$\begin{aligned}
n_i(t+\tau,x)=\;& n_i\left[1-\frac{\gamma_i\,(p-p(t,x+\Delta_x))_+}{2\overline{p}}-\frac{\gamma_i\,(p-p(t,x-\Delta_x))_+}{2\overline{p}}\right]\\
&+\left(n_i+\Delta_x\frac{\partial n_i}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 n_i}{\partial x^2}\right)\left[\frac{\gamma_i\,(p(t,x+\Delta_x)-p)_+}{2\overline{p}}\right]\\
&+\left(n_i-\Delta_x\frac{\partial n_i}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 n_i}{\partial x^2}\right)\left[\frac{\gamma_i\,(p(t,x-\Delta_x)-p)_+}{2\overline{p}}\right]\\
&+\tau G_i(p)n_i\left[1-\frac{\gamma_i\,(p-p(t,x+\Delta_x))_+}{2\overline{p}}-\frac{\gamma_i\,(p-p(t,x-\Delta_x))_+}{2\overline{p}}\right]\\
&+\tau G_i(p(t,x+\Delta_x))\left(n_i+\Delta_x\frac{\partial n_i}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 n_i}{\partial x^2}\right)\\
&\left[\frac{\gamma_i\,(p(t,x+\Delta_x)-p)_+}{2\overline{p}}\right]\\
&+\tau G_i(p(t,x-\Delta_x))\left(n_i-\Delta_x\frac{\partial n_i}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 n_i}{\partial x^2}\right)\\
&\left[\frac{\gamma_i\,(p(t,x-\Delta_x)-p)_+}{2\overline{p}}\right]+\mathcal{O}(\Delta_x^3).
\end{aligned}$$

Furthermore, rearranging terms yields

$$\begin{aligned}
n_i(t+\tau,x)=\;& n_i\left[1-\frac{\gamma_i\,(p-p(t,x+\Delta_x))_+}{2\overline{p}}-\frac{\gamma_i\,(p-p(t,x-\Delta_x))_+}{2\overline{p}}\right]\\
&+n_i\left[\frac{\gamma_i\,(p(t,x+\Delta_x)-p)_+}{2\overline{p}}+\frac{\gamma_i\,(p(t,x-\Delta_x)-p)_+}{2\overline{p}}\right]\\
&+\Delta_x\frac{\partial n_i}{\partial x}\left[\frac{\gamma_i\,(p(t,x+\Delta_x)-p)_+}{2\overline{p}}-\frac{\gamma_i\,(p(t,x-\Delta_x)-p)_+}{2\overline{p}}\right]\\
&+\frac{\Delta_x^2}{2}\frac{\partial^2 n_i}{\partial x^2}\left[\frac{\gamma_i\,(p(t,x+\Delta_x)-p)_+}{2\overline{p}}+\frac{\gamma_i\,(p(t,x-\Delta_x)-p)_+}{2\overline{p}}\right]\\
&+\tau G_i(p)n_i\left[1-\frac{\gamma_i\,(p-p(t,x+\Delta_x))_+}{2\overline{p}}-\frac{\gamma_i\,(p-p(t,x-\Delta_x))_+}{2\overline{p}}\right]\\
&+\tau G_i(p)n_i\left[\frac{\gamma_i\,(p(t,x+\Delta_x)-p)_+}{2\overline{p}}+\frac{\gamma_i\,(p(t,x-\Delta_x)-p)_+}{2\overline{p}}\right]
\end{aligned}$$

$$
\begin{aligned}
&+\Delta_x \frac{\partial n_i}{\partial x}\left[\tau G_i(p(t,x+\Delta_x))\frac{\gamma_i\,(p(t,x+\Delta_x)-p)_+}{2\overline{p}}\right]\\
&-\Delta_x \frac{\partial n_i}{\partial x}\left[\tau G_i(p(t,x-\Delta_x))\frac{\gamma_i\,(p(t,x-\Delta_x)-p)_+}{2\overline{p}}\right]\\
&+\frac{\Delta_x^2}{2}\frac{\partial^2 n_i}{\partial x^2}\left[\tau G_i(p(t,x+\Delta_x))\frac{\gamma_i\,(p(t,x+\Delta_x)-p)_+}{2\overline{p}}\right]\\
&+\frac{\Delta_x^2}{2}\frac{\partial^2 n_i}{\partial x^2}\left[\tau G_i(p(t,x-\Delta_x))\frac{\gamma_i\,(p(t,x-\Delta_x)-p)_+}{2\overline{p}}\right]+\mathcal{O}(\Delta_x^3).
\end{aligned}
$$

Then, using the property that $(f)_+ - (-f)_+ = (f)_+ - (f)_- = f$, we can simplify the latter equation to

$$
\begin{aligned}
n_i(t+\tau,x) = n_i&\left[1+\frac{\gamma_i\,(p(t,x+\Delta_x)-p)}{2\overline{p}}+\frac{\gamma_i\,(p(t,x-\Delta_x)-p)}{2\overline{p}}\right]\\
&+\Delta_x\frac{\partial n_i}{\partial x}\left[\frac{\gamma_i\,(p(t,x+\Delta_x)-p)_+}{2\overline{p}}-\frac{\gamma_i\,(p(t,x-\Delta_x)-p)_+}{2\overline{p}}\right]\\
&+\frac{\Delta_x^2}{2}\frac{\partial^2 n_i}{\partial x^2}\left[\frac{\gamma_i\,(p(t,x+\Delta_x)-p)_+}{2\overline{p}}+\frac{\gamma_i\,(p(t,x-\Delta_x)-p)_+}{2\overline{p}}\right]\\
&+\tau G_i(p)n_i\left[1+\frac{\gamma_i\,(p(t,x+\Delta_x)-p)}{2\overline{p}}+\frac{\gamma_i\,(p(t,x-\Delta_x)-p)}{2\overline{p}}\right]\\
&+\Delta_x\frac{\partial n_i}{\partial x}\left[\tau G_i(p(t,x+\Delta_x))\frac{\gamma_i\,(p(t,x+\Delta_x)-p)_+}{2\overline{p}}\right]\\
&-\Delta_x\frac{\partial n_i}{\partial x}\left[\tau G_i(p(t,x-\Delta_x))\frac{\gamma_i\,(p(t,x-\Delta_x)-p)_+}{2\overline{p}}\right]\\
&+\frac{\Delta_x^2}{2}\frac{\partial^2 n_i}{\partial x^2}\left[\tau G_i(p(t,x+\Delta_x))\frac{\gamma_i\,(p(t,x+\Delta_x)-p)_+}{2\overline{p}}\right]\\
&+\frac{\Delta_x^2}{2}\frac{\partial^2 n_i}{\partial x^2}\left[\tau G_i(p(t,x-\Delta_x))\frac{\gamma_i\,(p(t,x-\Delta_x)-p)_+}{2\overline{p}}\right]\\
&+\mathcal{O}(\Delta_x^3),
\end{aligned}
$$

from which, if $p$ is sufficiently regular, using that

$$
p(t,x\pm\Delta_x) = p \pm \Delta_x\frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}+\mathcal{O}(\Delta_x^3),
$$

we formally obtain

$$
n_i(t+\tau,x) = n_i\left[1+\frac{\gamma_i\left(\Delta_x\frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)}{2\overline{p}}+\frac{\gamma_i\left(-\Delta_x\frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)}{2\overline{p}}\right]
$$

$$
\begin{aligned}
&+\Delta_x \frac{\partial n_i}{\partial x}\left[\frac{\gamma_i\left(\Delta_x \frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)_+}{2\overline{p}}-\frac{\gamma_i\left(-\Delta_x \frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)_+}{2\overline{p}}\right]\\
&+\frac{\Delta_x^2}{2}\frac{\partial^2 n_i}{\partial x^2}\left[\frac{\gamma_i\left(\Delta_x \frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)_+}{2\overline{p}}+\frac{\gamma_i\left(-\Delta_x \frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)_+}{2\overline{p}}\right]\\
&+\tau G_i(p) n_i\left[1+\frac{\gamma_i\left(\Delta_x \frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)}{2\overline{p}}+\frac{\gamma_i\left(-\Delta_x \frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)}{2\overline{p}}\right]\\
&+\Delta_x \frac{\partial n_i}{\partial x}\left[\tau G_i\left(p+\Delta_x \frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)\frac{\gamma_i\left(\Delta_x \frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)_+}{2\overline{p}}\right]\\
&-\Delta_x \frac{\partial n_i}{\partial x}\left[\tau G_i\left(p-\Delta_x \frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)\frac{\gamma_i\left(-\Delta_x \frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)_+}{2\overline{p}}\right]\\
&+\frac{\Delta_x^2}{2}\frac{\partial^2 n_i}{\partial x^2}\left[\tau G_i\left(p+\Delta_x \frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)\frac{\gamma_i\left(\Delta_x \frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)_+}{2\overline{p}}\right]\\
&+\frac{\Delta_x^2}{2}\frac{\partial^2 n_i}{\partial x^2}\left[\tau G_i\left(p-\Delta_x \frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)\frac{\gamma_i\left(-\Delta_x \frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)_+}{2\overline{p}}\right]\\
&+\mathcal{O}(\Delta_x^3).
\end{aligned}
$$

Rearranging terms and neglecting terms of order higher than $\mathcal{O}(\Delta_x^3)$ we find

$$
\begin{aligned}
n_i(t+\tau,x) = n_i&\left[1+\frac{\gamma_i\Delta_x^2}{2\overline{p}}\frac{\partial^2 p}{\partial x^2}\right]+\Delta_x \frac{\partial n_i}{\partial x}\left[\frac{\gamma_i\left(\Delta_x \frac{\partial p}{\partial x}\right)_+}{2\overline{p}}-\frac{\gamma_i\left(-\Delta_x \frac{\partial p}{\partial x}\right)_+}{2\overline{p}}\right]\\
&+\tau G_i(p) n_i\left[1+\frac{\gamma_i\Delta_x^2}{2\overline{p}}\frac{\partial^2 p}{\partial x^2}\right]\\
&+\Delta_x \frac{\partial n_i}{\partial x}\left[\tau G_i\left(p+\Delta_x \frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)\frac{\gamma_i\left(\Delta_x \frac{\partial p}{\partial x}\right)_+}{2\overline{p}}\right]
\end{aligned}
$$

$$-\Delta_x \frac{\partial n_i}{\partial x}\left[\tau G_i\left(p-\Delta_x\frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)\frac{\gamma_i\left(-\Delta_x\frac{\partial p}{\partial x}\right)_+}{2\overline{p}}\right]+\mathcal{O}(\Delta_x^3).$$

Again, using that $(f)_+ - (-f)_+ = (f)_+ - (f)_- = f$, we can simplify the above equation to

$$\begin{aligned} n_i(t+\tau,x) = n_i\left[1+\frac{\gamma_i\Delta_x^2}{2\overline{p}}\frac{\partial^2 p}{\partial x^2}\right]+\frac{\gamma_i\Delta_x^2}{2\overline{p}}\frac{\partial n_i}{\partial x}\frac{\partial p}{\partial x}+\tau G_i(p)n_i\left[1+\frac{\gamma_i\Delta_x^2}{2\overline{p}}\frac{\partial^2 p}{\partial x^2}\right]\\ +\Delta_x\frac{\partial n_i}{\partial x}\left[\tau G_i\left(p+\Delta_x\frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)\frac{\gamma_i\left(\Delta_x\frac{\partial p}{\partial x}\right)_+}{2\overline{p}}\right]\\ -\Delta_x\frac{\partial n_i}{\partial x}\left[\tau G_i\left(p-\Delta_x\frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)\frac{\gamma_i\left(-\Delta_x\frac{\partial p}{\partial x}\right)_+}{2\overline{p}}\right]+\mathcal{O}(\Delta_x^3), \end{aligned}$$

from which, rearranging again terms, we obtain

$$\begin{aligned} n_i(t+\tau,x) = n_i+\frac{\gamma_i\Delta_x^2}{2\overline{p}}\left[n_i\frac{\partial^2 p}{\partial x^2}+\frac{\partial n_i}{\partial x}\frac{\partial p}{\partial x}\right]\\ +\tau G_i(p)n_i+\frac{\gamma_i\Delta_x^2\tau}{2\overline{p}}G_i(p)n_i\frac{\partial^2 p}{\partial x^2}\\ +\frac{\gamma_i\Delta_x^2\tau}{2\overline{p}}\frac{\partial n_i}{\partial x}\left[\alpha_i G\left(p+\Delta_x\frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)\left(\frac{\partial p}{\partial x}\right)_+\right.\\ \left.-G_i\left(p-\Delta_x\frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)\left(-\frac{\partial p}{\partial x}\right)_+\right]\\ +\mathcal{O}(\Delta_x^3). \end{aligned}$$

Now taking $n_i$ over to the left-hand side and dividing through by $\tau$ yields

$$\begin{aligned} \frac{n_i(t+\tau,x)-n_i}{\tau} = \frac{\gamma_i\Delta_x^2}{2\tau\overline{p}}\left[n_i\frac{\partial^2 p}{\partial x^2}+\frac{\partial n_i}{\partial x}\frac{\partial p}{\partial x}\right]\\ +G_i(p)n_i+\frac{\gamma_i\Delta_x^2}{2\overline{p}}G_i(p)n_i\frac{\partial^2 p}{\partial x^2}\\ +\frac{\gamma_i\Delta_x^2}{2\overline{p}}\frac{\partial n_i}{\partial x}\left[G_i\left(p+\Delta_x\frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)\left(\frac{\partial p}{\partial x}\right)_+\right.\\ \left.-G_i\left(p-\Delta_x\frac{\partial p}{\partial x}+\frac{\Delta_x^2}{2}\frac{\partial^2 p}{\partial x^2}\right)\left(-\frac{\partial p}{\partial x}\right)_+\right]\\ +\mathcal{O}(\Delta_x^3). \end{aligned}$$

Finally, letting $\Delta_x \to 0$ and $\tau \to 0$ in such a way that

$$\frac{\gamma_i \Delta_x^2}{2\tau \overline{p}} \to \mu_i, \quad \text{where} \quad \mu_i \in \mathbb{R}_+^*, \ i = 1, \dots, I,$$

we formally obtain the following PDEs

$$\frac{\partial n_i}{\partial t} = \mu_i \left( n_i \frac{\partial^2 p}{\partial x^2} + \frac{\partial n_i}{\partial x} \frac{\partial p}{\partial x} \right) + G_i(p) n_i, \quad i = 1, \dots, I, \quad (t, x) \in (0, \infty) \times \mathbb{R},$$

subject to the assumptions (13) on the mobility coefficients, $\mu_i$, which descend from assumptions (9) when conditions (11) hold. The above PDEs alongside the constitutive relation (A1) can then be rewritten as the continuum model (12).

## Appendix B: Addendum to *Step 8* of the Proof of Theorem 1

Differentiating the differential equation (54a) with respect to $z$ gives

$$-c \left(p'\right)' - \mu_1 \left[ p \left(p'\right)'' + 2\, p' \left(p'\right)' + p''\, p' \right] = \left[ \alpha_1 \frac{\mathrm{d}G}{\mathrm{d}p}\, p + \alpha_1\, G\left(p\right) \right] p',$$

and, given the conditions (54b), we complement the above differential equation with the following boundary conditions

$$p'(-\infty) = 0, \qquad p'(0^-) = -\frac{c}{\mu_1}.$$

On the other hand, differentiating the differential equation (54a) with respect to $c$ yields

$$\begin{aligned} &-c \left( \frac{\partial p}{\partial c} \right)' - \mu_1 \left[ p \left( \frac{\partial p}{\partial c} \right)'' + 2\, p' \left( \frac{\partial p}{\partial c} \right)' + p'' \frac{\partial p}{\partial c} \right] \\ &= \left[ \alpha_1 \frac{\mathrm{d}G}{\mathrm{d}p}\, p + \alpha_1\, G\left(p\right) \right] \frac{\partial p}{\partial c} + p' \end{aligned}$$

and, given the conditions (54b), we complement the above differential equation with the following boundary conditions

$$\frac{\partial p}{\partial c}(-\infty) = 0, \qquad \left( \frac{\partial p}{\partial c} \right)' (0^-) = -\frac{1}{\mu_1}.$$

Introducing the notation $f := p'$ and $g := \dfrac{\partial p}{\partial c}$, the above differential equations along with the corresponding boundary conditions can be rewritten, respectively, as

$$\begin{cases} -c\,(f)' - \left[a_1\,(f)'' + a_2\,(f)' + a_3\,f\right] = a_4\,f, \quad z \in (-\infty, 0), & \text{(B2a)} \\ f(-\infty) = 0, \qquad f(0^-) = -\dfrac{c}{\mu_1} & \text{(B2b)} \end{cases}$$

and

$$\begin{cases} -c\,(g)' - \left[a_1\,(g)'' + a_2\,(g)' + a_3\,g\right] = a_4\,g + f, \quad z \in (-\infty, 0), & \text{(B3a)} \\ g(-\infty) = 0, \qquad g'(0^-) = -\dfrac{1}{\mu_1}, & \text{(B3b)} \end{cases}$$

where the coefficients $a_1, \ldots, a_4$ are implicitly defined. Since $f < 0$ on $(-\infty, 0)$ and $g'(0^-) < 0$, noting that both $f(-\infty) = 0$ and $g(-\infty) = 0$ and the right-hand side of the differential equation (B3a) contains the additional term $f < 0$ compared to the right-hand side of the differential equation (B2a), we deduce that $g = \dfrac{\partial p}{\partial c} < 0$ on $(-\infty, 0)$, from which we conclude that $c \mapsto p(0^-)$ is monotonically decreasing.

## Appendix C: Numerical Scheme for the PDE Model (3)

To construct numerical solutions of the PDE system (3) we use a finite-volume scheme with an upwind stabilisation as well as a MUSCL reconstruction of interface values to increase the spatial order of accuracy. This scheme is built upon those we employed in Bubba et al. (2020); Lorenzi et al. (2023). The full details are provided below.

We first discretise the 1D domain into spatial cells of size $\Delta x$. The time discretisation is a forward Euler method. We allow the time-step to vary during the simulation using the CFL-like condition

$$\frac{\Delta t}{\Delta x} \max(\mu_1, \ldots, \mu_I) \max_x(|\nabla p|) < 1.$$

For each spatial cell interface, we reconstruct the values of the cell densities adopting a MUSCL approach. For the $i^{th}$ component of the solution we have

$$\begin{aligned} n^L_{i(j+1/2)} &= n^k_{i(j)} + 0.5\phi(r_j)(n^k_{i(j+1)} - n^k_{i(j)}), \\ n^R_{i(j+1/2)} &= n^k_{i(j)} - 0.5\phi(r_{j+1})(n^k_{i(j+2)} - n^k_{i(j+1)}), \\ n^L_{i(j-1/2)} &= n^k_{i(j-1)} + 0.5\phi(r_{j-1})(n^k_{i(j)} - n^k_{i(j-1)}), \\ n^R_{i(j-1/2)} &= n^k_{i(j)} - 0.5\phi(r_j)(n^k_{i(j+1)} - n^k_{i(j)}), \end{aligned}$$

where

$$r_j = \frac{n^k_{i(j)} - n^k_{i(j-1)}}{n^k_{i(j+1)} - n^k_{i(j)}},$$

and the function $\phi(\cdot)$ is the chosen flux-limiter. In particular, here we choose the *ospre* flux-limiter, i.e.

$$\phi(r) = \frac{1.5\left(r^2 + r\right)}{\left(r^2 + r + 1\right)}.$$

We then compute the gradient of the pressure at each cell interface. To do so, we first define

$$p^L_{j+1/2} = \sum_{i=1}^{I} \omega_i n^L_{i(j+1/2)}, \quad p^R_{j+1/2} = \sum_{i=1}^{I} \omega_i n^R_{i(j+1/2)}.$$

We then compute the quantities $\nabla p^L_{j+1/2}$ and $\nabla p^R_{j+1/2}$ using: a central differencing scheme in the interior of the spatial domain; a forward differencing scheme at the left endpoint of the spatial domain; and a backward differencing scheme at the right endpoint of the spatial domain. Finally, we employ the reconstruction

$$(\nabla p)_{j+1/2} = 0.5((\nabla p)^L_{j+1/2} + (\nabla p)^R_{j+1/2}).$$

The flux across the spatial cell interface for the discretisation of the $i^{th}$ component of the solution is given by

$$F^{n_i}_{j+1/2} = \mu_i \, n^L_{i(j+1/2)} \max(0, -(\nabla p)_{j+1/2}) + \mu_i \, n^R_{i(j+1/2)} \min(0, -(\nabla p)_{j+1/2}).$$

Altogether, upon splitting between conservative and non-conservative terms, the numerical scheme reads as

$$n^{k+1/2}_{i(j)} = n^k_{i(j)} + \Delta t \left[ -\frac{F^{n_i}_{j+1/2} - F^{n_i}_{j-1/2}}{\Delta x} \right],$$
$$n^{k+1}_{i(j)} = n^{k+1/2}_{i(j)} + \Delta t \, \alpha_i \, n^{k+1/2}_{i(j)} G(p^k_j).$$

**Acknowledgements** JAC was supported by the Advanced Grant Nonlocal-CPD (Nonlocal PDEs for Complex Particle Dynamics: Phase Transitions, Patterns and Synchronization) of the European Research Council Executive Agency (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 883363). JAC was also partially supported by the EPSRC grant number EP/V051121/1 and by the "Maria de Maeztu" Excellence Unit IMAG, reference CEX2020-001105-M, funded by MCIN/AEI/10.13039/501100011033/. TL gratefully acknowledges support from: the Italian Ministry of University and Research (MUR) through the grant PRIN 2020 project (No. 2020JLWP23) "Integrated Mathematical Approaches to Socio-Epidemiological Dynamics" (CUP: E15F21005420006)

and the grant PRIN2022-PNRR project (No. P2022Z7ZAJ) "A Unitary Mathematical Framework for Modelling Muscular Dystrophies" (CUP: E53D23018070001) funded by the European Union - Next Generation EU; and the Istituto Nazionale di Alta Matematica (INdAM) and the Gruppo Nazionale per la Fisica Matematica (GNFM). TL would also like to thank Luigi Preziosi for insightful discussions during the development of the project.

**Data Availability** The code used for numerical simulations and data for this manuscript is available upon request.



## References

Alasio L, Bruna M, Fagioli S, Schulz S (2022) Existence and regularity for a system of porous medium equations with small cross-diffusion and nonlocal drifts. Nonlinear Anal 223:113064

Ambrosi D, Preziosi L (2002) On the closure of mass balance models for tumor growth. Math Models Methods Appl Sci 12(05):737–754

Baker EL, Lu J, Yu D, Bonnecaze RT, Zaman MH (2010) Cancer cell stiffness: integrated roles of three-dimensional matrix stiffness and transforming potential. Biophys J 99(7):2048–2057

Baker RE, Parker A, Simpson MJ (2019) A free boundary model of epithelial dynamics. J Theor Biol 481:61–74

Basan M, Risler T, Joanny JF, Sastre-Garau X, Prost J (2009) Homeostatic competition drives tumor growth and metastasis nucleation. HFSP J 3(4):265–272

Batlle E, Wilkinson DG (2012) Molecular mechanisms of cell segregation and boundary formation in development and tumorigenesis. Cold Spring Harb Perspect Biol 4(1):008227

Berendsen J, Burger M, Pietschmann J-F (2017) On a cross-diffusion model for multiple species with nonlocal interaction and size exclusion. Nonlinear Anal 159:10–39

Bertsch M, Gurtin ME, Hilhorst D, Peletier LA (1985) On interacting populations that disperse to avoid crowding: preservation of segregation. J Math Biol 23:1–13

Bertsch M, Gurtin ME, Hilhorst D et al (1987) On interacting populations that disperse to avoid crowding: the case of equal dispersal velocities. Nonlinear Anal 11(4):493–499

Bertsch M, Gurtin ME, Hilhorst D (1987) On a degenerate diffusion equation of the form $c(z)t = \vartheta(zx)x$ with application to population dynamics. J Differ Equ 67(1):56–89

Bertsch M, Dal Passo R, Mimura M (2010) A free boundary problem arising in a simplified tumour growth model of contact inhibition. Interfaces Free Bound 12(2):235–250

Bertsch M, Hilhorst D, Izuhara H, Mimura M et al (2012) A nonlinear parabolic-hyperbolic system for contact inhibition of cell-growth. Differ Equ Appl 4(1):137–157

Bertsch M, Hilhorst D, Izuhara H, Mimura M, Wakasa T (2015) Travelling wave solutions of a parabolic-hyperbolic system for contact inhibition of cell-growth. Eur J Appl Math 26(3):297–323

Bubba F, Perthame B, Pouchol C, Schmidtchen M (2020) Hele-shaw limit for a system of two reaction-(cross-) diffusion equations for living tissues. Arch Ration Mech Anal 236(2):735–766

Bubba F, Lorenzi T, Macfarlane FR (2020) From a discrete model of chemotaxis with volume-filling to a generalized Patlak-Keller-Segel model. Proc R Soc A 476(2237):20190871

Burger M, Carrillo JA, Pietschmann JF, Schmidtchen M (2020) Segregation and gap formation in cross-diffusion models. Interfaces Free Bound 22(2):175–203

Burger M, Carrillo JA, Pietschmann J-F, Schmidtchen M (2020) Segregation effects and gap formation in cross-diffusion models. Interfaces Free Bound 22(2):175–203

Byrne HM, Chaplain M (1997) Free boundary value problems associated with the growth and development of multicellular spheroids. Eur J Appl Math 8(6):639–658
Byrne H, Drasdo D (2009) Individual-based and continuum models of growing cell populations: a comparison. J Math Biol 58:657–687
Byrne H, Preziosi L (2003) Modelling solid tumour growth using the theory of mixtures. Math Med Biol 20(4):341–366
Carrillo JA, Huang Y, Schmidtchen M (2018) Zoology of a nonlocal cross-diffusion model for two species. SIAM J Appl Math 78(2):1078–1104
Carrillo JA, Fagioli S, Santambrogio F, Schmidtchen M (2018) Splitting schemes and segregation in reaction cross-diffusion systems. SIAM J Math Anal 50(5):5695–5718
Carrillo JA, Murakawa H, Sato M, Togashi H, Trush O (2019) A population dynamics model of cell-cell adhesion incorporating population pressure and density saturation. J Theor Biol 474:14–24
Celora GL, Byrne HM, Zois CE, Kevrekidis PG (2021) Phenotypic variation modulates the growth dynamics and response to radiotherapy of solid tumours under normoxia and hypoxia. J Theor Biol 527:110792
Chaplain MAJ, Graziano L, Preziosi L (2006) Mathematical modelling of the loss of tissue compression responsiveness and its role in solid tumour development. Math Med Biol 23(3):197–229. https://doi.org/10.1093/imammb/dql009
Chaplain M, Giverso C, Lorenzi T, Preziosi L (2019) Derivation and application of effective interface conditions for continuum mechanical models of cell invasion through thin membranes. SIAM J Appl Math 79(5):2011–2031
Chaplain M, Lorenzi T, Macfarlane FR (2020) Bridging the gap between individual-based and continuum models of growing cell populations. J Math Biol 80(1):343–371
Charras G, Sahai E (2014) Physical influences of the extracellular environment on cell migration. Nat Rev Mol Cell Biol 15(12):813–824
Chen B-J, Wu J-S, Tang Y-J, Tang Y-L, Liang X-H (2020) What makes leader cells arise: intrinsic properties and support from neighboring cells. J Cell Physiol 235(12):8983–8995
Ciarletta P, Foret L, Ben Amar M (2011) The radial growth phase of malignant melanoma: multi-phase modelling, numerical simulations and linear stability analysis. J R Soc Interface 8(56):345–368
Ciavolella G, Perthame B (2021) Existence of a global weak solution for a reaction-diffusion problem with membrane conditions. J Evol Equ 21(2):1513–1540
Ciavolella G, David N, Poulain A (2024) Effective interface conditions for a porous medium type problem. Interfaces Free Bound 26(2):161–188
Crossley RM, Painter KJ, Lorenzi T, Maini PK, Baker RE (2024) Phenotypic switching mechanisms determine the structure of cell migration into extracellular matrix under the 'go-or-grow' hypothesis. Math Biosci, 109240
David N (2023) Phenotypic heterogeneity in a model of tumour growth: existence of solutions and incompressible limit. Comm Partial Differ Equ 48(4):678–710
David N, Dębiec T, Mandal M, Schmidtchen M (2024) A degenerate cross-diffusion system as the inviscid limit of a nonlocal tissue growth model. SIAM J Math Anal 56(2):2090–2114
Drasdo D (2005) Coarse graining in simulated cell populations. Adv Complex Syst 8(02n03):319–363
Drasdo D, Hoehme S (2012) Modeling the impact of granular embedding media, and pulling versus pushing cells on growing cell clones. New J Phys 14(5):055025
Dyson L, Maini PK, Baker RE (2012) Macroscopic limits of individual-based models for motile cell populations with volume exclusion. Phys Rev E 86(3):031903
Fadai NT, Simpson MJ (2020) New travelling wave solutions of the porous-fisher model with a moving boundary. J Phys A: Math Theor 53(9):095601
Falcó C, Baker RE, Carrillo JA (2024) A local continuum model of cell-cell adhesion. SIAM J Appl Math 84(3):17–42
Fozard JA, Byrne HM, Jensen OE, King JR (2010) Continuum approximations of individual-based models for epithelial monolayers. Math Med Biol 27(1):39–74
Fraser S, Keynes R, Lumsden A (1990) Segmentation in the chick embryo hindbrain is defined by cell lineage restrictions. Nature 344(6265):431–435
Galiano G, Selgas V (2014) On a cross-diffusion segregation problem arising from a model of interacting particles. Nonlinear Anal Real World Appl 18:34–49
Girardin L, Hilhorst D (2022) Spatial segregation limit of traveling wave solutions for a fully nonlinear strongly coupled competitive system. Electron Res Arch 30(5):1748–1773

Giverso C, Lorenzi T, Preziosi L (2022) Effective interface conditions for continuum mechanical models describing the invasion of multiple cell populations through thin membranes. Appl Math Lett 125:107708
Gurtin ME, Pipkin AC (1984) A note on interacting populations that disperse to avoid crowding. Q J Appl Math 42(1):87–94
Gwiazda P, Perthame B, Świerczewska-Gwiazda A (2019) A two-species hyperbolic-parabolic model of tissue growth. Commun Partial Differ Equ 44(12):1605–1618
Han YL, Pegoraro AF, Li H, Li K, Yuan Y, Xu G, Gu Z, Sun J, Hao Y, Gupta SK et al (2020) Cell swelling, softening and invasion in a three-dimensional breast cancer model. Nat Phys 16(1):101–108
Huang Y, Hong W, Wei X (2022) The molecular mechanisms and therapeutic strategies of EMT in tumor progression and metastasis. J Hematol Oncol 15(1):129
Hughes BD (1996) Random walks and random environments. Oxford University Press, UK
Irvine KD, Rauskolb C (2001) Boundaries in development: formation and function. Annu Rev Cell Dev Biol 17(1):189–214
Johnston ST, Simpson MJ, Baker RE (2012) Mean-field descriptions of collective migration with strong adhesion. Phys Rev E 85(5):051922
Kalukula Y, Stephens AD, Lammerding J, Gabriele S (2022) Mechanics and functional consequences of nuclear deformations. Nat Rev Mol Cell Biol 23(9):583–602
Konen J, Summerbell E, Dwivedi B, Galior K, Hou Y, Rusnak L, Chen A, Saltz J, Zhou W, Boise LH et al (2017) Image-guided genomics of phenotypically heterogeneous populations reveals vascular signalling during symbiotic collective cancer invasion. Nat Commun 8(1):1–15
Lamouille S, Xu J, Derynck R (2014) Molecular mechanisms of epithelial-mesenchymal transition. Nat Rev Mol Cell Biol 15(3):178–196
Langenberg T, Brand M (2005) Lineage restriction maintains a stable organizer cell population at the zebrafish midbrain-hindbrain boundary. Development 132:3209–3216
Lieberman MA, Glaser L (1981) Density-dependent regulation of cell growth: an example of a cell-cell recognition phenomenon. J Membr Biol 63:1–11
Lorenzi T, Painter KJ (2022) Trade-offs between chemotaxis and proliferation shape the phenotypic structuring of invading waves. Int J Non-Linear Mech 139:103885
Lorenzi T, Lorz A, Perthame B (2017) On interfaces between cell populations with different mobilities. Kinet Relat Models 10(1):299–311
Lorenzi T, Murray PJ, Ptashnyk M (2020) From individual-based mechanical models of multicellular systems to free-boundary problems. Interfaces Free Bound 22(2):205–244
Lorenzi T, Macfarlane FR, Painter KJ (2023) Derivation and travelling wave analysis of phenotype-structured haptotaxis models of cancer invasion. Eur J Appl Math, pp 1–33
Lorenzi T, Painter KJ, Villa C (2024) Phenotype structuring in collective cell migration: a tutorial of mathematical models and methods. arXiv preprint arXiv:2410.13629
Lorenzi T, Perthame B, Ruan X (2021) Invasion fronts and adaptive dynamics in a model for the growth of cell populations with heterogeneous mobility. Eur J Appl Math, 1–18
Lv J, Liu Y, Cheng F, Li J, Zhou Y, Zhang T, Zhou N, Li C, Wang Z, Ma L et al (2021) Cell softness regulates tumorigenicity and stemness of cancer cells. EMBO J 40(2):106123
Macfarlane FR, Chaplain M, Lorenzi T (2020) A hybrid discrete-continuum approach to model turing pattern formation. Math Biosci Eng 17(6):7442–7479
Macfarlane FR, Ruan X, Lorenzi T, Ruan X, Lorenzi T (2022) Individual-based and continuum models of phenotypically heterogeneous growing cell populations. AIMS Bioeng 9(1):68–92
Macfarlane FR, Lorenzi T, Painter KJ (2022) The impact of phenotypic heterogeneity on chemotactic self-organisation. Bull Math Biol 84(12):143
Masaeli M, Gupta D, O'Byrne S, Tse H, Gossett DR, Tseng P, Utada AS, Jung HJ, Young S, Clark AT et al (2016) Multiparameter mechanical and morphometric screening of cells. Sci Rep 6(1):37863
Mok S, Al Habyan S, Ledoux C, Lee W, MacDonald K, McCaffrey L, Moraes C (2020) Mapping cellular-scale internal stiffness in 3D tissues with smart material hydrogel probes. Nat Commun 11:4757
Murphy RJ, Buenzli PR, Baker RE, Simpson MJ (2020) Mechanical cell competition in heterogeneous epithelial tissues. Bull Math Biol 82(10):130
Murphy RJ, Buenzli PR, Baker RE, Simpson MJ (2021) Travelling waves in a free boundary mechanobiological model of an epithelial tissue. Appl Math Lett 111:106636
Oelschläger K (1989) On the derivation of reaction-diffusion equations as limit dynamics of systems of moderately interacting stochastic processes. Probab Theory Relat Fields 82(4):565–586

Penington CJ, Hughes BD, Landman KA (2011) Building macroscale models from microscale probabilistic models: a general probabilistic approach for nonlinear diffusion and multispecies phenomena. Phys Rev E 84(4):041120
Perthame B, Quirós F, Vázquez JL (2014) The hele-shaw asymptotics for mechanical models of tumor growth. Arch Ration Mech Anal 212:93–127
Pillay S, Byrne HM, Maini PK (2017) Modeling angiogenesis: a discrete to continuum description. Phys Rev E 95(1):012410
Pillay S, Byrne HM, Maini PK (2018) The impact of exclusion processes on angiogenesis models. J Math Biol 77:1721–1759
Preziosi L, Tosin A (2009) Multiphase modelling of tumour growth and extracellular matrix interaction: mathematical tools and applications. J Math Biol 58:625–656
Ranft J, Basan M, Elgeti J, Joanny JF, Prost J, Jülicher F (2010) Fluidization of tissues by cell division and apoptosis. Proc Natl Acad Sci 107(49):20863–20868
Rianna C, Radmacher M, Kumar S (2020) Direct evidence that tumor cells soften when navigating confined spaces. Mol Biol Cell 31(16):1726–1734
Roose T, Chapman SJ, Maini PK (2007) Mathematical models of avascular tumor growth. SIAM Rev 49(2):179–208
Simpson MJ, Merrifield A, Landman KA, Hughes BD (2007) Simulating invasion with cellular automata: connecting cell-scale and population-scale properties. Phys Rev E 76(2):021918
Simpson MJ, Baker RE, McCue SW (2011) Models of collective cell spreading with variable cell aspect ratio: a motivation for degenerate diffusion models. Phys Rev E 83(2):021901
Simpson MJ, Murphy KM, McCue SW, Buenzli PR (2024) Discrete and continuous mathematical models of sharp-fronted collective cell migration and invasion. R Soc Open Sci 11(5):240126
Swaminathan V, Mythreye K, O'Brien ET, Berchuck A, Blobe GC, Superfine R (2011) Mechanical stiffness grades metastatic potential in patient tumor cells and in cancer cell lines. Cancer Res 71(15):5075–5080
Tambyah TA, Murphy RJ, Buenzli PR, Simpson MJ (2020) A free boundary mechanobiological model of epithelial tissues. Proc R Soc A 476(2243):20200528
Vilchez Mercedes SA, Bocci F, Levine H, Onuchic JN, Jolly MK, Wong PK (2021) Decoding leader cells in collective cancer invasion. Nat Rev Cancer 21(9):592–604
Wang X-C, Tang Y-L, Liang X-H (2023) Tumour follower cells: a novel driver of leader cells in collective invasion. Int J Oncol 63(4):1–17
Zanotelli MR, Zhang J, Reinhart-King CA (2021) Mechanoresponsive metabolism in cancer cell migration and metastasis. Cell Metab 33(7):1307–1321
Zeltser LM, Larsen CW, Lumsden A (2001) A new developmental compartment in the forebrain regulated by lunatic fringe. Nat Neurosci 4(7):683–684
Zills G, Datta T, Malmi-Kakkada AN (2023) Enhanced mechanical heterogeneity of cell collectives due to temporal fluctuations in cell elasticity. Phys Rev E 107(1):014401