Inverse problem approaches for mutation laws in heterogeneous tumours with local and nonlocal dynamics

Alwuthaynani, Maher; Eftimie, Raluca; Trucu, Dumitru

Published in:
Mathematical Biosciences and Engineering

DOI:
10.3934/mbe.2022171

Publication date:
2022

Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/mbe

MBE, 19(4): 3720–3747.
DOI: 10.3934/mbe.2022171
Received: 20 October 2021
Revised: 05 December 2021
Accepted: 22 December 2021
Published: 10 February 2022

Research article

Inverse problem approaches for mutation laws in heterogeneous tumours with local and nonlocal dynamics

Maher Alwuthaynani[1], Raluca Eftimie[2] and Dumitru Trucu[1,*]

[1] Division of Mathematics, University of Dundee, Dundee DD1 4HN, Scotland, UK

[2] Laboratoire Mathématiques de Besançon, UMR-CNRS 6623, Université de Bourgogne Franche-Comté, 16 Route de Gray, Besançon 25000, France

* **Correspondence:** Email: trucu@maths.dundee.ac.uk.

Abstract: Cancer cell mutations occur when cells undergo multiple cell divisions, and these mutations can be spontaneous or environmentally-induced. The mechanisms that promote and sustain these mutations are still not fully understood. This study deals with the identification (or reconstruction) of the usually unknown cancer cell mutation law, which lead to the transformation of a primary tumour cell population into a secondary, more aggressive cell population. We focus on local and nonlocal mathematical models for cell dynamics and movement, and identify these mutation laws from macroscopic tumour snapshot data collected at some later stage in the tumour evolution. In a local cancer invasion model, we first reconstruct the mutation law when we assume that the mutations depend only on the surrounding cancer cells (i.e., the ECM plays no role in mutations). Second, we assume that the mutations depend on the ECM only, and we reconstruct the mutation law in this case. Third, we reconstruct the mutation when we assume that there is no prior knowledge about the mutations. Finally, for the nonlocal cancer invasion model, we reconstruct the mutation law that depends on the cancer cells and on the ECM. For these numerical reconstructions, our approximations are based on the finite difference method combined with the finite elements method. As the inverse problem is ill-posed, we use the Tikhonov regularisation technique in order to regularise the solution. Stability of the solution is examined by adding additive noise into the measurements.

Keywords: inverse problems; mutation identification; Tikhonov regularisation; tumour growth

1. Introduction

The beginning of a primary solid tumour is the result of a single normal cell that is transformed as a result of mutations in certain key genes. Cells can mutate spontaneously, or mutations can be environmentally induced. Mutations occur during cell division, and most of the time the immune

system can recognise mutated cells and eliminate them. When the immune system fails to eliminate cells with mutations in genes that control cell proliferation, cells become cancerous. It is known that all cancer cell lines have at least one mutation, with most cancer cell lines having more than one mutation; e.g., in [1] 137 oncogenic mutations were identified in 14 out of 24 known cancer genes in 60 human cancer cell lines. Moreover, cancer cells are genetically unstable and cells inside the solid tumours keep mutating leading to very heterogeneous tumour masses.

The mechanisms behind the mutation pressure are still not fully understood. Experimental studies have shown that some changes in the extracellular matrix (ECM) can correlate with sustained cell proliferative signalling and an increased risk of developing cancer [2]. Other studies have shown that culturing cells for long times in stiff hydrogels can lead to the subclonal selection of genomic aberrations in cells [3], thus suggesting that the ECM properties could impact the mutation status of cells in solid tumours. Other studies suggested that the maintenance of cells at high density in the absence of proliferation leads to an increase in mutagenesis following cell division [4]. Therefore, there seem to be different mechanisms that can trigger and influence the mutation rate of cells.

Mutated cells not only incur sustained proliferation (see Figure 1), but can also exhibit migrational and invasion properties [5], eventually leading to cancer metastasis. The invasion of surrounding tissue is the result of ECM degradation and remodelling by the cancer cells (which can secrete various proteolytic enzymes, such as matrix metalloproteinases (MMPs)) as well as other cells in the microenvironment. The last decades have seen the development of numerous theoretical studies based on mathematical models, which investigate computationally the biological mechanisms behind cancer cell invasion into the tissue. Most of these models are single-scale models; see [6–11] and references therein. More recently, multi-scale mathematical models have started to be developed to consider also the multi-scale aspects of various biological processes occurring during cancer invasion [12–16].

The main issue faced by all these single-scale and especially multi-scale mathematical models for cancer growth and invasion is parameter estimation. The last few years have seen the publication of various mathematical studies that try to estimate numerically different model parameters using inverse problem formulations [17–21]. These studies focus on identifying the magnitude of isotropic [21] and anisotropic tumour diffusion [18], the magnitude of tumour growth rate [19, 21], the strength/location of tumour-induced tissue deformation [19], the position of the blood vessels that act as a source for the oxygen concentration that influences tumour growth [17], the location of the source of tumours [20].

To our knowledge, no studies have tried to estimate the mutation laws for cells inside heterogeneous tumours. To address this aspect, in this current study we consider a simplified problem with only two cancer cell populations (one mutating into the second one), that can exhibit both random and directed haptotactic movement. Moreover, we assume that the cells can degrade and remodel the surrounding ECM density. Furthermore, the type of interactions among cancer cells and between cells and ECM is not always very clear: some simpler mathematical models for cancer growth and invasion consider local cell-cell and cell-ECM interactions [8, 9], while other more complex models consider non-local cell-cell and cell-ECM interactions [11,16,22,23]. In this study, we consider both type of models, local and non-local, and estimate the mutation laws for both cases. Finally, we assume here the knowledge of additional information in terms of both exact and noisy measurements of the tumour constituent density at some later time in the tumour evolution. We test our inversion approach on several cancer proliferation laws that are usually used in cancer modelling: logistic and Gompertz proliferation.

This paper is structured as follows. In Section 2 we describe the local mathematical model for the

dynamics of the two cancer cell sub-populations and the extracellular matrix (ECM). In Sections 3 and 3.1 we formulate the inverse problem for this local model under the assumptions of logistic cell proliferation and mutation depending only on cancer cell density. In Section 3.4 we formulate the inverse problem for this local model under the assumption of logistic cell proliferation and mutation depending on both cell and ECM density. In Section 4 we describe a non-local model for two cancer cell populations and their interactions with the ECM, while in Section A we present the numerical approach for this forward nonlocal model. In Section 4.1 we formulate the inverse problem for this nonlocal model for cancer invasion. We summarise and discuss our results in Section 5.

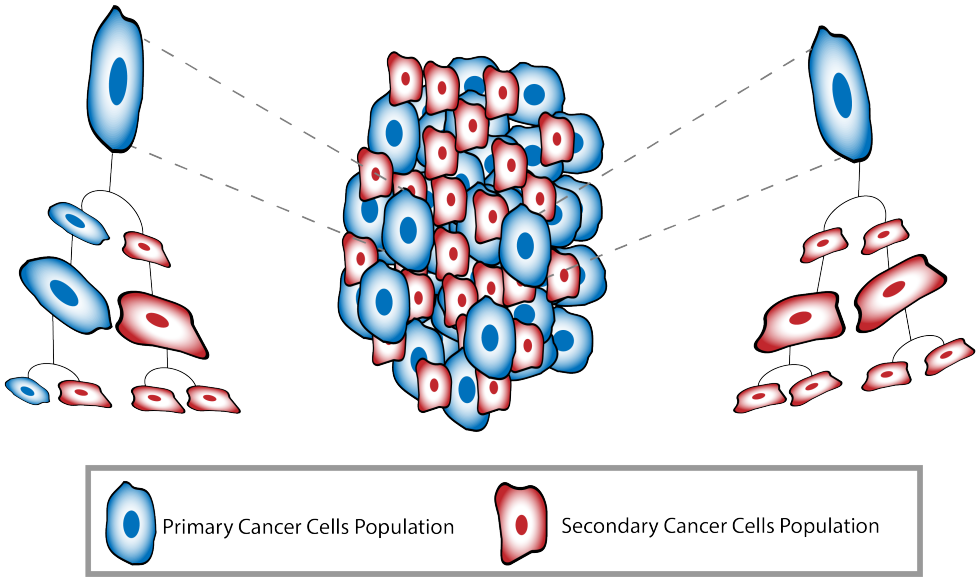

Figure 1. Schematic of cancer cells proliferation.

2. Mathematical model for two local cancer cell sub-populations

In this study we consider two populations of tumour cells: a primary cell population $c_1(x,t)$ that can mutate into a more aggressive secondary cell population $c_2(x,t)$. These cell populations exercise a spatial redistribution via random movement (with diffusion coefficients D_1 and D_2) and directed movement towards extracellular matrix (ECM) gradients $v(x,t)$ (with haptotactic coefficients η_1 and η_2). Moreover, the two populations undergo logistic growth (at the same rate μ_c), up to a carrying capacity (K_c). We also assume that the ECM undergoes degradation (at a rate ρ) and remodelling (at a rate μ_v). The above assumptions are described by the equations:

$$\frac{\partial c_1}{\partial t} = \underbrace{D_1\Delta c_1}_{\text{diffusion}} - \underbrace{\eta_1\nabla\cdot(c_1\nabla v)}_{\text{haptotactic movement}} + \underbrace{\mu_c c_1\left(1-\frac{c_1+c_2+v}{K_c}\right)}_{\text{logistic proliferation}} - \underbrace{\omega(t)}_{\substack{\text{mutation}\\\text{switch}}}\underbrace{Q(\cdot,\cdot)}_{\substack{\text{unknown}\\\text{mutation}}}, \tag{2.1a}$$

$$\frac{\partial c_2}{\partial t} = \underbrace{D_2\Delta c_2}_{\text{diffusion}} - \underbrace{\eta_2\nabla\cdot(c_2\nabla v)}_{\text{haptotactic movement}} + \underbrace{\mu_c c_2\left(1-\frac{c_1+c_2+v}{K_c}\right)}_{\text{logistic proliferation}} + \underbrace{\omega(t)}_{\substack{\text{mutation}\\\text{switch}}}\underbrace{Q(\cdot,\cdot)}_{\substack{\text{unknown}\\\text{mutation}}}, \tag{2.1b}$$

$$\frac{\partial v}{\partial t} = \underbrace{-\rho\,(c_1 + c_2)\,v}_{\text{degradation}} + \underbrace{\mu_v\,(K_c - v - c_1 - c_2)^+}_{\text{remodelling term}}. \tag{2.1c}$$

where $(K_c - v - c_1 - c_2)^+ := max\{(K_c - v - c_1 - c_2), 0\}$. Finally, the unknown term $Q(\cdot,\cdot)$ represents the mutation law of cell subpopulation c_1 into cell subpopulation c_2, which is assumed to be mediated by a time-dependant mutation enhancement $\omega(t)$ that is known *a priori* and is taken here of the form

$$\omega(t) := \frac{\left(1 + tanh\left(\frac{t-t_{1,2}}{t_s}\right)\right)}{2},$$

where $t_{1,2}$ is the time at which mutations from c_1 to c_2 start occurring, and $t_s > 0$ is a time-steepness coefficient for this mutation law.

The mutation law $Q(\cdot,\cdot)$ is considerd here unknown due to either unknown dependance on the primary cell population c_1, or unknown dependance on the ECM v, or unknown dependance on both primary tumour cell population and ECM. In this study we investigate three assumptions related to this mutation term, namely: (i) mutation depends linearly on the density of primary tumour cells as experimental studies have shown that tumour hypoxia, generated by high tumour cell density, is linearly correlated with an increase in genomic changes toward more aggressive tumours [24]; (ii) mutation depends linearly on the density of primary tumour cells, and nonlinearly on the ECM density [3]; (iii) mutation law is very general and depends autonomously on both the primary tumour and ECM, as suggested by various experimental studies regarding the role of the tumour microenvironment in cancer cells mutation process [3, 24, 25]. Thus mathematically, these cases correspond to three inverse problems that seek to identify the unknown mutation law $Q(\cdot,\cdot)$ in the following three situations:

(i) Mutation depends linearly on the density of primary tumour cell sub-population c_1 but does not depend at all on ECM, and so this is given by the unknown term

$$Q(c_1, v) := \widetilde{Q}_1(c_1) = \delta_0 c_1, \tag{2.2}$$

with δ_0 representing the unknown mutation rate.

(ii) Mutation depends in a known linear manner on c_1 and in an unknown nonlinear way on the density of ECM. The unknown dependence on v is denoted mathematically by the unknown function $\widetilde{Q}_2(v)$, and so the entire mutation law is therefore of the form

$$Q(c_1, v) := \delta_0 c_1 \widetilde{Q}_2(v), \tag{2.3}$$

with $\delta_0 > 0$ here being considered known. A usual choice for $\widetilde{Q}_2(v)$ is of the form [26, 27]:

$$\widetilde{Q}_2(v) := \begin{cases} \frac{exp\left(\frac{-1}{\kappa^2-(1-v(x,t))^2}\right)}{exp\left(-\frac{1}{\kappa^2}\right)} & \text{if} \quad 1-\kappa < v\,(x,t) < 1, \\ 0, & \text{otherwise,} \end{cases} \tag{2.4}$$

where $\kappa > 0$ is a certain level of ECM beyond which mutations can occur.

(iii) Mutation is given by an unknown general nonlinear law $Q(c_1, v)$ that is autonomous in c_1 and v, which will be reconstructed from the data available at a specific later time in the tumour evolution.

To complete the description of the model, we mention that the initial conditions for the cancer-ECM dynamics described by model (2.1) are:

$$c_1(x,0) := c_{1,0}(x), \quad c_2(x,0) := c_{2,0}(x) \quad \text{and} \quad v(x,0) := v_0(x), \quad \text{for } x \in \Omega. \tag{2.5}$$

Here, $c_{1,0}(\cdot)$, $c_{2,0}(\cdot)$ and $v_0(\cdot)$ give the initial distributions of the primary cell subpopulation, mutated cell subpopulation and ECM, respectively. Furthermore, we assume that the cells and the ECM components do not leave the tissue region Ω, and therefore we consider zero Neumann boundary conditions:

$$\left.\frac{\partial c_1}{\partial n}\right|_{\partial\Omega} = 0, \quad \left.\frac{\partial c_2}{\partial n}\right|_{\partial\Omega} = 0 \quad \text{and} \quad \left.\frac{\partial v}{\partial n}\right|_{\partial\Omega} = 0, \tag{2.6}$$

where $n(\xi)$ is the usual normal direction at any given tissue boundary point $\xi \in \partial\Omega$.

Throughout the following sections, we refer to the tumour dynamics model (2.1) together with the initial and boundary conditions (2.5) and (2.6) as the *"forward model"*.

3. Inverse problem formulation for the unknown cancer cell mutation

We start with the *forward model* defined by the tumour dynamics model (2.1) in the presence of the initial and boundary conditions (2.5) and (2.6). Our goal is to reconstruct the unknown cancer cells mutation law $\mathcal{Q}(\cdot)$ from additional information given by measurements of the cancer cells and ECM densities taken at some later time $t_f := T > 0$ in the tumour evolution. These measurements are therefore given in the form of the following functions on Ω, which are considered to be known:

$$c_1^*(\cdot) : \Omega \to \mathbb{R} \quad \text{for the cancer subpopulation } c_1, \tag{3.1a}$$
$$c_2^*(\cdot) : \Omega \to \mathbb{R} \quad \text{for the cancer subpopulation } c_2, \tag{3.1b}$$
$$v^*(\cdot) : \Omega \to \mathbb{R} \quad \text{for the ECM density.} \tag{3.1c}$$

In the following, we will explore the reconstruction of the unknown cancer mutation law $\mathcal{Q}(\cdot)$ when the known measurements $c_1^*(x)$, $c_2^*(x)$ and $v^*(x)$ will be given both as exact (accurate) data and as noisy data, $\forall x \in \Omega$.

3.1. Inverse problem setup: forward solver for the retrieval of mutation laws in Cases (i) and (ii)

In this section we outline in a unitary manner the forward solver involved in the retrieval of the mutation laws corresponding to Cases (i) and (ii) that require the retrieval of $\widetilde{\mathcal{Q}}_1(c_1)$ and $\widetilde{\mathcal{Q}}_2(v)$, respectively. To that end, for $r = 1, 2$, denoting by e_r either the primary tumour or the ECM, *i.e.*, $e_r \in \{c_1, v\}$, enables us to proceed with addressing simultaneously both cases by simply referring to the retrieval of the term compactly denoted as $\widetilde{\mathcal{Q}}_r(e_r)$ that is specified by

$$e_r := \begin{cases} c_1, & if \quad r = 1, \\ v, & if \quad r = 2, \end{cases} \quad \text{and subsequently} \quad \widetilde{\mathcal{Q}}_r(e_r) := \begin{cases} \widetilde{\mathcal{Q}}_1(c_1), & if \quad r = 1, \\ \widetilde{\mathcal{Q}}_2(v), & if \quad r = 2. \end{cases} \tag{3.2}$$

We start by considering a uniform discretisation $\mathcal{G}_\Omega := \{(x_i, y_j)\}_{i,j=1\ldots N}$ of step size $\Delta x = \Delta y > 0$ for a square maximal tissue region $\Omega \subset \mathbb{R}^2$ where the tumour exercises its dynamics. At any given time

$t \in [0, t_f]$ the discretisations of cancer cells densities $c_1(\cdot, t)$ and $c_2(\cdot, t)$ as well as the density of ECM $v(\cdot, t)$ are therefore given by the $N \times N$ matrices $\tilde{c}_1(t) := \{\tilde{c}_{1,i,j}(t)\}_{i,j=1\ldots N}$, $\tilde{c}_2(t) := \{\tilde{c}_{2,i,j}(t)\}_{i,j=1\ldots N}$ and $\tilde{v}(t) := \{\tilde{v}_{i,j}(t)\}_{i,j=1\ldots N}$, with $\tilde{c}_{1,i,j}(t) := c_1((x_i, y_j), t)$, $\tilde{c}_{2,i,j}(t) := c_2((x_i, y_j), t)$ and $\tilde{v}_{i,j}(t) := v((x_i, y_j), t)$, $\forall\ i, j = 1 \ldots N$. Correspondingly, in the following $\tilde{e}_r(t)$ will denote either $\tilde{c}_1(t)$ or $\tilde{v}(t)$, *i.e.*, $\tilde{e}_r(t) \in \{\tilde{c}_1(t), \tilde{v}(t)\}$, as required by model (3.2).

Throughout this study we assume that we have *a priori* knowledge that the cumulated ECM and cancer densities do not exceed the tissue carrying capacity K_c. Under this assumption, the unknown mutation law can be written in terms of an unknown (for the moment) function $m^{c_1^*, c_2^*, v^*} : [0, K_c] \to [0, \infty)$. Moreover, this unknown function $m^{c_1^*, c_2^*, v^*}$ will be appropriately identified within a suitable family of functions $\mathfrak{M}^1$ such that the corresponding solution for the tumour model (2.1) will match the measurements given in model (3.1). Thus, denoting by $\widetilde{Q}_r^{c_1^*, c_2^*, v^*}(\cdot)$ the unknown mutation term for which the corresponding solution of model (2.1) matches measurement model (3.1), at each (x_i, y_j) we can write this as

$$\widetilde{Q}_r^{c_1^*, c_2^*, v^*}(\tilde{e}_{r,i,j}^{m^{c_1^*, c_2^*, v^*}}(t)) := \mathcal{M}_{i,j}^1(\tilde{e}_r^{m^{c_1^*, c_2^*, v^*}}(t), m^{c_1^*, c_2^*, v^*}),$$

where $\mathcal{M}^1(\cdot, \cdot) := \{\mathcal{M}_{i,j}^1(\cdot, \cdot)\}_{i,j=1\ldots N}$, with $\mathcal{M}^1(\cdot, \cdot) : \mathbb{R}^{N\times N} \times \mathfrak{M}^1 \to \mathbb{R}^{N\times N}$ representing a *"trial mutation operator"* that will be specified below alongside the family of functions $\mathfrak{M}^1$. Furthermore, $\tilde{e}_r^{m^{c_1^*, c_2^*, v^*}}(t) := \{\tilde{e}_{r,i,j}^{m^{c_1^*, c_2^*, v^*}}(t)\}_{i,j=1..N}$ represents the solution for the density of either the primary cell population (if $r = 1$) or the ECM (if $r = 2$) that is obtained for model (2.1) when, instead of the unknown term $\widetilde{Q}_r(e_r)$, in the mutation law we use the trial mutation term $\mathcal{M}^1(\cdot, m^{c_1^*, c_2^*, v^*})$.

Next, we consider an uniform discretisation for the domain $[0, K_c]$ that is given by an equally spaced grid $\mathcal{G}_M^1 := \{\eta_l\}_{l=1\ldots M}$ of step size $\Delta\eta > 0$. On this discretised domain, the unknown function $m^{c_1^*, c_2^*, v^*}$ is identified through a suitable approximation within the following $M-$dimensional space of functions associated with $\mathcal{G}_M^1$, namely

$$\begin{aligned}\mathfrak{M}^1 := &\left\{ m : [0, K_c] \to \mathbb{R} \;\middle|\; m_{|_{E_l}} = \sum_{p=0,1} m(\eta_{l+p})\phi_{l+p}, \quad \forall E_l \in \mathcal{G}_M^{1,tiles} \right\} \\ &\text{with the family of intervals } \mathcal{G}_M^{1,tiles} := \{E_l := [\eta_l, \eta_{l+1}] \mid l = 1 \ldots M-1\}, \\ &\text{and } \forall\ E_l \in \mathcal{G}_M^{1,tiles},\ \{\phi_{l+p}\}_{p=0,1} \text{ describe the usual linear shape functions on } E_l.\end{aligned} \tag{3.3}$$

Thus, for any candidate function $m \in \mathfrak{M}^1$, the corresponding *trial mutation operator* $\mathcal{M}^1$ has each of its components $\mathcal{M}_{i,j}$, $\forall\ i, j = 1 \ldots N$, given by

$$\begin{aligned}&\mathcal{M}_{i,j}^1(\tilde{e}_r^m(t), m) := m_{|_{E_l}}(\tilde{e}_{r,i,j}^m(t)), \\ &\text{with index } l \text{ being independent of its choice within the associated set of indices } \Lambda_{i,j}, \text{ namely:} \\ &\Lambda_{i,j} := \{l \in \{1, \ldots, M-1\} \mid \exists E_l \in \mathcal{G}_M^{1,tiles} \text{ such that } \tilde{e}_{r,i,j}^m(t) \in E_l\}.\end{aligned} \tag{3.4}$$

Here, as per model (3.2), $\tilde{e}_r^m(t) := \{\tilde{e}_{ri,j}^m(t)\}_{i,j=1\ldots N}$ represents either the solutions for the density of primary cancer cell population, $\{\tilde{c}_{1,i,j}^m(t)\}_{i,j=1\ldots N}$, or for the density of ECM, $\{\tilde{v}_{i,j}^m(t)\}_{i,j=1\ldots N}$, which alongside the density of mutated cell population, $\{\tilde{c}_{1,i,j}^m(t)\}_{i,j=1\ldots N}$, are obtained with model (2.1) when this uses within the mutation law the trial mutation term $\mathcal{M}^1(\cdot, m) := \{\mathcal{M}_{i,j}^1(\cdot, m)\}_{i,j=1\ldots N}$ given in model (3.4) instead of the unknown term $\widetilde{Q}_r(\cdot)$. Finally, the *trial mutation form* for the full mutation law

given in models (2.2) and (2.3) for Cases (*i*) and (*ii*), respectively, is denoted by $\overline{\mathcal{M}}_r(\tilde{c}_1^m(t), \tilde{v}^m(t), m) := \{\overline{\mathcal{M}}_{r,i,j}(\tilde{c}_1^m(t), \tilde{v}^m(t), m)\}_{i.j=1\ldots N}$ and is given by:

$$\overline{\mathcal{M}}_{r,i,j}(\tilde{c}_1^m(t), \tilde{v}^m(t), m) := \begin{cases} \mathcal{M}_{i,j}^1(\tilde{c}_1^m(t), m), & if \quad r = 1, \\ \delta_0 \tilde{c}_{1,i,j}^m(t) \mathcal{M}_{i,j}^1(\tilde{v}^m(t), m), & if \quad r = 2. \end{cases} \tag{3.5}$$

Therefore, in space-discretised form, model (3.4) that uses $\mathcal{M}^1(\cdot, m)$ can therefore be written as

$$\frac{\partial}{\partial t} \begin{bmatrix} \tilde{c}_1^m \\ \tilde{c}_2^m \\ \tilde{v}^m \end{bmatrix} = \begin{bmatrix} \mathcal{H}^1(\tilde{c}_1^m, c_2^m, \tilde{v}^m, m) \\ \mathcal{H}^2(\tilde{c}_1^m, c_2^m, \tilde{v}^m, m) \\ \mathcal{H}^3(\tilde{c}_1^m, c_2^m, \tilde{v}^m) \end{bmatrix}, \tag{3.6}$$

Here, $\mathcal{H}^1(\cdot,\cdot,\cdot,\cdot) = \{\mathcal{H}_{i,j}^1(\cdot,\cdot,\cdot,\cdot)\}_{i,j=1\ldots N}$ represents the spatial discretisation corresponding to Eq (2.1a). Each of its components $\mathcal{H}_{i,j}^1(\cdot,\cdot,\cdot,\cdot)$, $\forall\ i, j = 1 \ldots N$, are given by

$$\begin{aligned} &\mathcal{H}_{i,j}^1\left(\tilde{c}_1^m(t), \tilde{c}_2^m(t), \tilde{v}^m(t), m\right) := \\ &\quad \frac{D_1}{(\Delta x)^2}\left(\tilde{c}_{1,i-1,j}^m(t) + \tilde{c}_{1,i+1,j}^m(t) + \tilde{c}_{1,i,j-1}^m(t) + \tilde{c}_{1,i,j+1}^m(t) - 4\tilde{c}_{1,i,j}^m(t)\right) \\ &\quad - \frac{\eta_1}{2(\Delta x)^2}\left(\left(\tilde{c}_{1,i,j}^m(t) + \tilde{c}_{1,i+1,j}^m(t)\right)\left(\tilde{v}_{i+1,j}^m(t) - \tilde{v}_{i,j}^m(t)\right) - \left(\tilde{c}_{1,i,j}^m(t) + \tilde{c}_{1,i-1,j}^m(t)\right)\left(\tilde{v}_{i,j}^m(t) - \tilde{v}_{i-1,j}^m(t)\right)\right. \\ &\quad \left. + \left(\tilde{c}_{1,i,j}^m(t) + \tilde{c}_{1,i,j+1}^m(t)\right)\left(\tilde{v}_{i,j+1}^m(t) - \tilde{v}_{i,j}^m(t)\right) - \left(\tilde{c}_{1,i,j}^m(t) + \tilde{c}_{1,i,j-1}^m(t)\right)\left(\tilde{v}_{i,j}^m(t) - \tilde{v}_{i,j-1}^m(t)\right)\right) \\ &\quad + \mu_c \tilde{c}_{1,i,j}^m(t)\left(K_c - \tilde{c}_{1,i,j}^m(t) - \tilde{c}_{2,i,j}^m(t) - \tilde{v}_{i,j}^m(t)\right) - \omega(t)\overline{\mathcal{M}}_{r,i,j}(\tilde{c}_1^m(t), \tilde{v}^m(t), m). \end{aligned} \tag{3.7}$$

Similarly, $\mathcal{H}^2(\cdot,\cdot,\cdot,\cdot) = \{\mathcal{H}_{i,j}^2(\cdot,\cdot,\cdot,\cdot)\}_{i,j=1\ldots N}$ represents the spatial discretisation corresponding to Eq (2.1b), and each of its components $\mathcal{H}_{i,j}^2(\cdot,\cdot,\cdot,\cdot)$, $\forall\ i, j = 1 \ldots N$, are given by

$$\begin{aligned} &\mathcal{H}_{i,j}^2(\tilde{c}_2^m(t), \tilde{c}_2^m(t), \tilde{v}^m(t), m) := \\ &\quad \frac{D_2}{(\Delta x)^2}\left(\tilde{c}_{2,i-1,j}^m(t) + \tilde{c}_{2,i+1,j}^m(t) + \tilde{c}_{2,i,j-1}^m(t) + \tilde{c}_{2,i,j+1}^m(t) - 4\tilde{c}_{2,i,j}^m(t)\right) \\ &\quad - \frac{\eta_2}{2(\Delta x)^2}\left(\left(\tilde{c}_{2,i,j}^m(t) + \tilde{c}_{2,i+1,j}^m(t)\right)\left(\tilde{v}_{i+1,j}^m(t) - \tilde{v}_{i,j}^m(t)\right) - \left(\tilde{c}_{2,i,j}^m(t) + \tilde{c}_{2,i-1,j}^m(t)\right)\left(\tilde{v}_{i,j}^m(t) - \tilde{v}_{i-1,j}^m(t)\right)\right. \\ &\quad \left. + \left(\tilde{c}_{2,i,j}^m(t) + \tilde{c}_{2,i,j+1}^m(t)\right)\left(\tilde{v}_{i,j+1}^m(t) - \tilde{v}_{i,j}^m(t)\right) - \left(\tilde{c}_{2,i,j}^m(t) + \tilde{c}_{2,i,j-1}^m(t)\right)\left(\tilde{v}_{i,j}^m(t) - \tilde{v}_{i,j-1}^m(t)\right)\right) \\ &\quad + \mu_c \tilde{c}_{2,i,j}^m(t)(K_c - \tilde{c}_{1,i,j}^m(t) - \tilde{c}_{2,i,j}^m(t) - \tilde{v}_{i,j}^m(t)) + \omega(t)\overline{\mathcal{M}}_{r,i,j}(\tilde{c}_1^m(t), \tilde{v}^m(t), m). \end{aligned} \tag{3.8}$$

Finally, $\mathcal{H}^3(\cdot,\cdot,\cdot) = \{\mathcal{H}_{i,j}^3(\cdot,\cdot,\cdot)\}_{i,j=1\ldots N}$ represents the discretisation of the ECM Eq (2.1c), and each of its components $\mathcal{H}_{i,j}^3(\cdot,\cdot)$, $\forall\ i, j = 1 \ldots N$, are given by

$$\mathcal{H}_{i,j}^3(\tilde{c}_1^m(t), \tilde{c}_2^m(t), \tilde{v}^m(t)) := -\rho(\tilde{c}_{1,i,j}^m(t) + \tilde{c}_{2,i,j}^m(t))\tilde{v}_{i,j}^m(t) + \mu_v(K_c - \tilde{c}_{1,i,j}^m(t) - \tilde{c}_{2,i,j}^m(t) - \tilde{v}_{i,j}^m(t))^+. \tag{3.9}$$

Consider now a time discretisation $\{t_n\}_{n=0\ldots L}$ with time step $\Delta t := T/L$. For each $n \in \{0, \ldots, L\}$, a

simple Euler time-marching scheme can be written for model (3.6) via the associated operator

$$K_m : \mathbb{R}^{N\times N} \times \mathbb{R}^{N\times N} \times \mathbb{R}^{N\times N} \to \mathbb{R}^{N\times N} \times \mathbb{R}^{N\times N} \times \mathbb{R}^{N\times N}$$

given by

$$K_m\left(\begin{bmatrix}\tilde{c}_1^{m,n}\\ \tilde{c}_2^{m,n}\\ \tilde{v}^{m,n}\end{bmatrix}\right) := \begin{bmatrix}\tilde{c}_1^{m,n}\\ \tilde{c}_2^{m,n}\\ \tilde{v}^{m,n}\end{bmatrix} + \Delta t \begin{bmatrix}\mathcal{H}^1(\tilde{c}_1^{m,n}, \tilde{c}_2^{m,n}, \tilde{v}^{m,n}, m)\\ \mathcal{H}^2(\tilde{c}_1^{m,n}, \tilde{c}_2^{m,n}, \tilde{v}^{m,n}, m)\\ \mathcal{H}^3(\tilde{c}_1^{m,n}, \tilde{c}_2^{m,n}, \tilde{v}^{m,n})\end{bmatrix}, \tag{3.10}$$

where $\tilde{c}_1^{m,n} := \tilde{c}_1^m(t_n)$, $\tilde{c}_2^{m,n} := \tilde{c}_2^m(t_n)$ and $\tilde{v}^{m,n} := \tilde{v}^m(t_n)$. The right-hand-side operators are given by

$$\begin{aligned}\mathcal{H}^1(\tilde{c}_1^{m,n}, \tilde{c}_2^{m,n}, \tilde{v}^{m,n}, m) &:= \mathcal{H}^1(\tilde{c}_1^m(t_n), \tilde{c}_2^m(t_n), \tilde{v}^m(t_n), m),\\ \mathcal{H}^2(\tilde{c}_1^{m,n}, \tilde{c}_2^{m,n}, \tilde{v}^{m,n}, m) &:= \mathcal{H}^2(\tilde{c}_1^m(t_n), \tilde{c}_2^m(t_n), \tilde{v}^m(t_n), m),\\ \mathcal{H}^3(\tilde{c}_1^{m,n}, \tilde{c}_2^{m,n}, \tilde{v}^{m,n}) &:= \mathcal{H}^3(\tilde{c}_1^m(t_n), \tilde{c}_2^m(t_n), \tilde{v}^m(t_n)).\end{aligned}$$

This allows us to formulate the *"forward operator" K* between the family of functions $\mathfrak{M}^1$ (where we search for the appropriate cancer cells mutation function $m^{c_1^*,c_2^*,v^*}$) and the space where the discretised measurements model (3.1) are recorded. Hence, the *"forward operator" K* is defined as

$$K : \mathcal{S} \to \mathbb{R}^{N\times N} \times \mathbb{R}^{N\times N} \times \mathbb{R}^{N\times N}$$

given by

$$K(m) := \underbrace{K_m \circ K_m \circ \cdots\cdots \circ K_m}_{L \text{ times}}\left(\begin{bmatrix}\tilde{c}_{1,0}\\ \tilde{c}_{2,0}\\ \tilde{v}_0\end{bmatrix}\right) \tag{3.11}$$

where $\tilde{c}_{1,0} := \{c_{1,0}(x_i, y_j)\}_{i,j=1,\ldots,N}$, $\tilde{c}_{2,0} := \{c_{2,0}(x_i, y_j)\}_{i,j=1,\ldots,N}$ and $\tilde{v}_0 := \{v_0(x_i, y_j)\}_{i,j=1,\ldots,N}$ are the discretised initial conditions (2.5) for the governing tumour *forward model* (2.1). Hence, for each $m \in \mathfrak{M}^1$, the forward operator K gives the spatio-temporal progression of the initial condition $[\tilde{c}_{1,0}, \tilde{c}_{2,0}, \tilde{v}_0]^T$ under the invasion model (2.1), which is obtained when the cell mutation law at each instance of time $t > 0$ involves the corresponding trial mutation operator $\mathcal{M}(\cdot, m)$ instead of the unknown mutation terms $\widetilde{Q}_r(\cdot)$.

3.2. The inverse problem regularisation approach for mutation laws in Cases (i) and (ii)

From models (3.10) and (3.11) we have that our forward operator K is given as a finite composition of affine functions of the form

$$\mathfrak{M}^1 \ni m \longmapsto K_m \in \boldsymbol{\ell}^2(\boldsymbol{\ell}^2(\mathcal{E}\times\mathcal{E}\times\mathcal{E}); \boldsymbol{\ell}^2(\mathcal{E}\times\mathcal{E}\times\mathcal{E})). \tag{3.12}$$

Here, $\boldsymbol{\ell}^2(\boldsymbol{\ell}^2(\mathcal{E}\times\mathcal{E}\times\mathcal{E}); \boldsymbol{\ell}^2(\mathcal{E}\times\mathcal{E}\times\mathcal{E}))$ is the usual finite-dimensional Bochner space of square integrable vector-value functions [28] with respect to the counting measure (see [29]) that are defined on $\boldsymbol{\ell}^2(\mathcal{E}\times\mathcal{E}\times\mathcal{E})$ and take values in $\boldsymbol{\ell}^2(\mathcal{E}\times\mathcal{E}\times\mathcal{E})$, and $\mathcal{E} := \{E_{i,j}\}_{i,j=1\ldots N}$ represents the standard basis of elementary matrices associated with the grid $\mathcal{G}_\Omega$. As a direct consequence, we immediately obtain that this operator is both continuous and compact, from where we obtain that K is also closed and

sequentially bounded [28]. Therefore, K satisfies the hypotheses assumed in [30] that ensure convergence for the nonlinear Tikhonov regularisation strategy given by the functionals $\{J_\alpha\}_{\alpha>0}$,

$$J_\alpha : \mathfrak{M}^1 \to \mathbb{R}, \quad \forall \alpha > 0,$$

defined by

$$J_\alpha(m) := \left\| K(m) - \begin{bmatrix} \tilde{c}_1^* \\ \tilde{c}_2^* \\ \tilde{v}^* \end{bmatrix} \right\|_2^2 + \alpha \|m\|_2^2, \quad \forall m \in \mathfrak{M}^1. \tag{3.13}$$

The minimisation of these functionals enable us to identify $m^{c_1^*, c_2^*, v^*}$ as the limit $\alpha \to 0$ of the points of minimum m^α of J_α (these points correspond to the smallest discrepancy between the data measurements and the solution of our system that uses m^α as a mutation law). The two norms involved in model (3.13) represent the usual Euclidean norms on the corresponding finite dimensional spaces. Indeed, while the first is the standard Euclidean norm on $\mathbb{R}^{N\times N} \times \mathbb{R}^{N\times N} \times \mathbb{R}^{N\times N}$, the second is also the Euclidean norm induced on the $M-$dimensional space of functions $\mathfrak{M}^1$ via the standard isomorphism that we establish between $\mathfrak{M}^1$ and $\mathbb{R}^M$ by which each $m \in \mathfrak{M}^1$ is uniquely represented through its nodal values $\{m(\eta_l)\}_{l=1\ldots M}$ with respect to the linear basis functions $\{\bar{\phi}_l\}_{l=1\ldots M}$ associated to $\mathcal{G}_M$ [31]:

$$\text{since } m = \sum_{l=1\ldots M} m(\eta_l)\bar{\phi}_l, \quad \text{we therefore make the identification: } \; m \equiv \{m(\eta_l)\}_{l=1\ldots M}. \tag{3.14}$$

Finally, in model (3.13), $\tilde{c}_1^*$, $\tilde{c}_2^*$ and $\tilde{v}^*$ represent the discretised measurements of the densities of cancer cells and ECM given in Eqs (3.1a)–(3.1c), i.e., $\tilde{c}_1^* := \{c_1^*(x_i, y_j)\}_{i,j=1,\ldots,N}$, $\tilde{c}_2^* := \{c_2^*(x_i, y_j)\}_{i,j=1,\ldots,N}$ and $\tilde{v}^* := \{v^*(x_i, y_j)\}_{i,j=1,\ldots,N}$. We assume that these data measurements are either exact or are corrupted by a certain noise level $\delta \geq 0$. Thus, maintaining for simplicity the measurements notation unchanged, these measurements are given by

$$\tilde{c}_1^*(x) = \tilde{c}_{1exact}^*(x) + \delta\gamma_{c_1}(x), \tag{3.15a}$$
$$\tilde{c}_2^*(x) = \tilde{c}_{2exact}^*(x) + \delta\gamma_{c_2}(x), \tag{3.15b}$$
$$\tilde{v}^*(x) = \tilde{v}_{exact}^*(x) + \delta\gamma_v(x), \tag{3.15c}$$

where, $\forall\ x \in \Omega$, we have that $\tilde{c}_{1exact}^*(x)$, $\tilde{c}_{2exact}^*(x)$ and $\tilde{v}_{exact}^*(x)$ describe the exact data, and $\gamma_{c_1}(x)$, $\gamma_{c_2}(x)$ and $\gamma_v(x)$ are signal-independent noise generated from a Gaussian normal distribution with mean zero and standard deviations σ_{c_1}, σ_{c_2} and σ_v, respectively, given by

$$\begin{cases} \sigma_{c_1} := \frac{1}{\lambda(\Omega)} \int\limits_\Omega \tilde{c}_{1exact}^*(x)\, dx, \\ \sigma_{c_2} := \frac{1}{\lambda(\Omega)} \int\limits_\Omega \tilde{c}_{2exact}^*(x)\, dx, \\ \sigma_v := \frac{1}{\lambda(\Omega)} \int\limits_\Omega \tilde{v}_{exact}^*(x)\, dx, \end{cases} \tag{3.16}$$

with $\lambda(\cdot)$ being the usual Lebesgue measure. In the numerical results below, we generate the random variables $\gamma_{c_1}(x)$, $\gamma_{c_2}(x)$ and $\gamma_v(x)$ via MATLAB function *normrnd* by taking $\{\gamma_{c_1}(x_i, y_j)\}_{i,j=1\ldots N} := normrnd(0, \sigma_{c_1}, N \times N)$, $\{\gamma_{c_2}(x_i, y_j)\}_{i,j=1\ldots N} := normrnd(0, \sigma_{c_2}, N \times N)$ and $\{\gamma_v(x_i, y_j)\}_{i,j=1\ldots N} := normrnd(0, \sigma_v, N \times N)$.

3.3. Numerical reconstruction of the unknown mutation laws terms in Cases (i) and (ii)

We explore now the inversion approach that we formulated so far in the context of the forward model (2.1) by proceeding with the reconstruction of the unknown terms involved in the mutation laws in Cases (i) and (ii), namely $\widetilde{Q}_1(c_1)$ and $\widetilde{Q}_2(v)$, respectively.

Initial Conditions The initial conditions (2.5) that we consider in the computations for the forward model (2.1) are as follows:

$$c_{1,0}(x) := 0.5\left(exp\left(-\frac{\|x-(2,2)\|_2^2}{0.03}\right) - exp\,(-9.407)\right), \tag{3.17a}$$

$$c_{2,0}(x) := 0, \tag{3.17b}$$

$$v_0(x) := 0.5 + 0.3 \cdot sin\,(4\pi \cdot \|x\|_2)\,, \qquad \forall\, x \in \Omega. \tag{3.17c}$$

Here, we assume that $c_{2,0}(x) = 0$ because this second cell population will arise after a period of time following mutations of the first cell population c_1.

To identify the cancer cells mutation law, we consider both exact and noisy measurement data model (3.15) as additional information for the forward model (2.1) in the presence of initial conditions (3.17) and boundary conditions (2.6). Specifically, in each of the two cases, we assume that the exact data (namely $\tilde{c}^*_{1,exact}(x)$, $\tilde{c}^*_{2,exact}(x)$ and $\tilde{v}^*_{exact}(x)$) that appear in model (3.15) are given by the solutions densities for primary tumour cells population, $\bar{c}_1(x,t)$, mutated cells population, $\bar{c}_2(x,t)$, and ECM, and $\bar{v}(x,t)$, evaluated at the final time $t_f > 0$, *i.e.*,

$$\tilde{c}^*_{1,exact}(x) := \bar{c}_1(x,t_f), \quad \tilde{c}^*_{2,exact}(x) := \bar{c}_2(x,t_f) \quad \text{and} \quad \tilde{v}^*_{exact}(x) := \bar{v}(x,t_f), \quad \forall x \in \Omega, \tag{3.18}$$

which are obtained from the forward model (2.1) as follows:

- for Case (*i*): we assume that the mutation law is of the form given in model (2.2) but when parameter $\delta_0 > 0$ is considered known and has the value given in parameter Table 1;
- for Case (*ii*): we assume that the mutation law is of the form given by model (2.3) with the known term $\widetilde{Q}_2(v)$ specified in model (2.4).

For each regularisation parameter $\alpha > 0$ considered here, the minimisation process for J_α is initiated with $m_0 = \mathbb{I} \times 10^{-3}$, (where $\mathbb{I}$ represents the M vector of ones), and for the actual minimisation we employed here the nonlinear minimisation MATLAB function *lsqnonlin*. Finally, since there are no data to test the trial mutation operators beyond the *maximal accessible region* $\mathcal{A}_c$ and $\mathcal{A}_v$ defined by the minimum and maximum values of the solution, i.e.,

$$\mathcal{A}_c := [\bar{c}_1^{min}, \bar{c}_1^{max}], \text{ with: } \quad \bar{c}_1^{min} := \min_{(x,t)\in\Omega\times[0,T]} c_1(x,t), \qquad \bar{c}_1^{max} := \max_{(x,t)\in\Omega\times[0,T]} c_1(x,t), \tag{3.19}$$

$$\mathcal{A}_v := [\bar{v}^{min}, \bar{v}^{max}], \text{ with: } \quad \bar{v}^{min} := \min_{(x,t)\in\Omega\times[0,T]} v(x,t), \qquad \bar{v}^{max} := \max_{(x,t)\in\Omega\times[0,T]} v(x,t), \tag{3.20}$$

the reconstructions in this section will be attempted only for the restriction of the sought mutation laws to $\mathcal{A}_c$ and $\mathcal{A}_v$. We need to emphasise that $\mathcal{A}_c$ is used to reconstruct $\widetilde{Q}_1(c_1)$, while $\mathcal{A}_v$ is used to reconstruct $\widetilde{Q}_2(v)$.

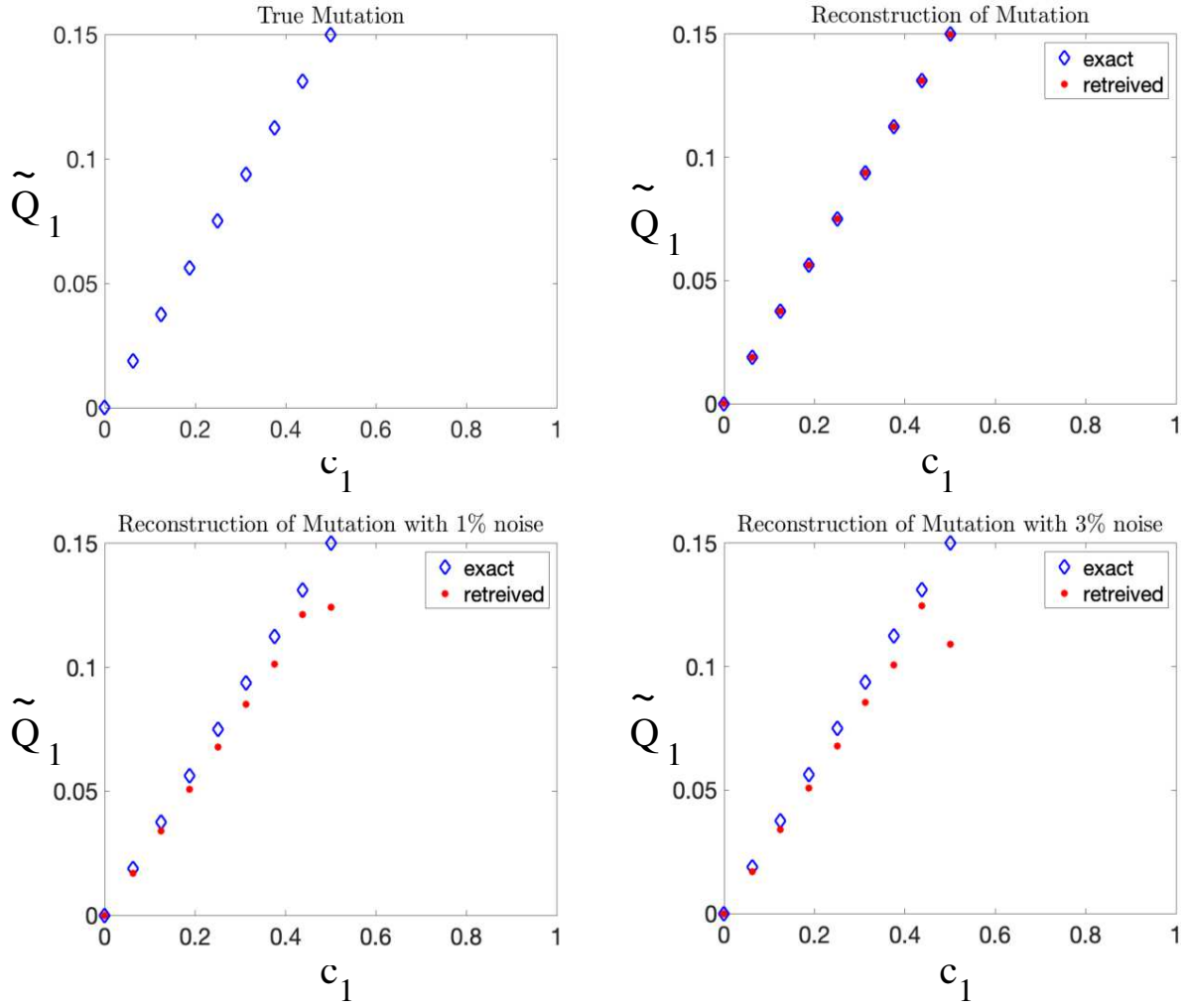

Figure 2. Reconstruction of the unknown mutation law term $\widetilde{Q}_1(c_1)$ that appears in mutation law given in model (2.2) for model (2.1). First row: (left) the true mutation law, and (right) the reconstructed mutation law restricted to $\mathcal{A}_c$ with no noise and $\alpha^* = 10^{-12}$. Second row: (left) the reconstructed mutation law obtained for 1% noisy data and $\alpha^* = 10^{-5}$; and (right) the reconstructed mutation law obtained for 3% noisy data and $\alpha^* = 10^{-12}$. For all plots in this figure: 1) the first axis represents the the values for $c_1 \in [\bar{c}_1^{min}, \bar{c}_1^{max}]$; 2) second axis represents the values for mutation. The numerical simulations are obtained using the parameters given in Table 1.

An acceptable numerical reconstruction of the mutation law $m^{c_1^*,c_2^*,v^*}$, i.e.,

$$m^{c_1^*,c_2^*,v^*} := m^{\alpha^*}, \tag{3.21}$$

is obtained for the choice of the regularisation parameter α^*, which throughout this work is selected based on a standard discrepancy principle argument [32].

Figure 2 shows the reconstruction of the cancer cell mutation law $\widetilde{Q}_1(c_1)$ for model (2.1) in the presence of the measurements given by models (3.15) and (3.18) that are considered here both exact

and affected by a level of noise $\delta \in \{1\%, 3\%\}$. The first row of this figure shows from left to right the true mutation law restricted at the maximal accessible region $\mathcal{A}_c$ (where the reconstruction is being attempted) and the reconstruction of the mutation law equation $\widetilde{Q}_1(c_1)$ given by model (2.2) on $\mathcal{A}_c$ with no noise, respectively. The second row of the figure show from left to right the reconstruction of the mutation law on $\mathcal{A}_c$ with 1%, and 3% of noise in the measured data, respectively.

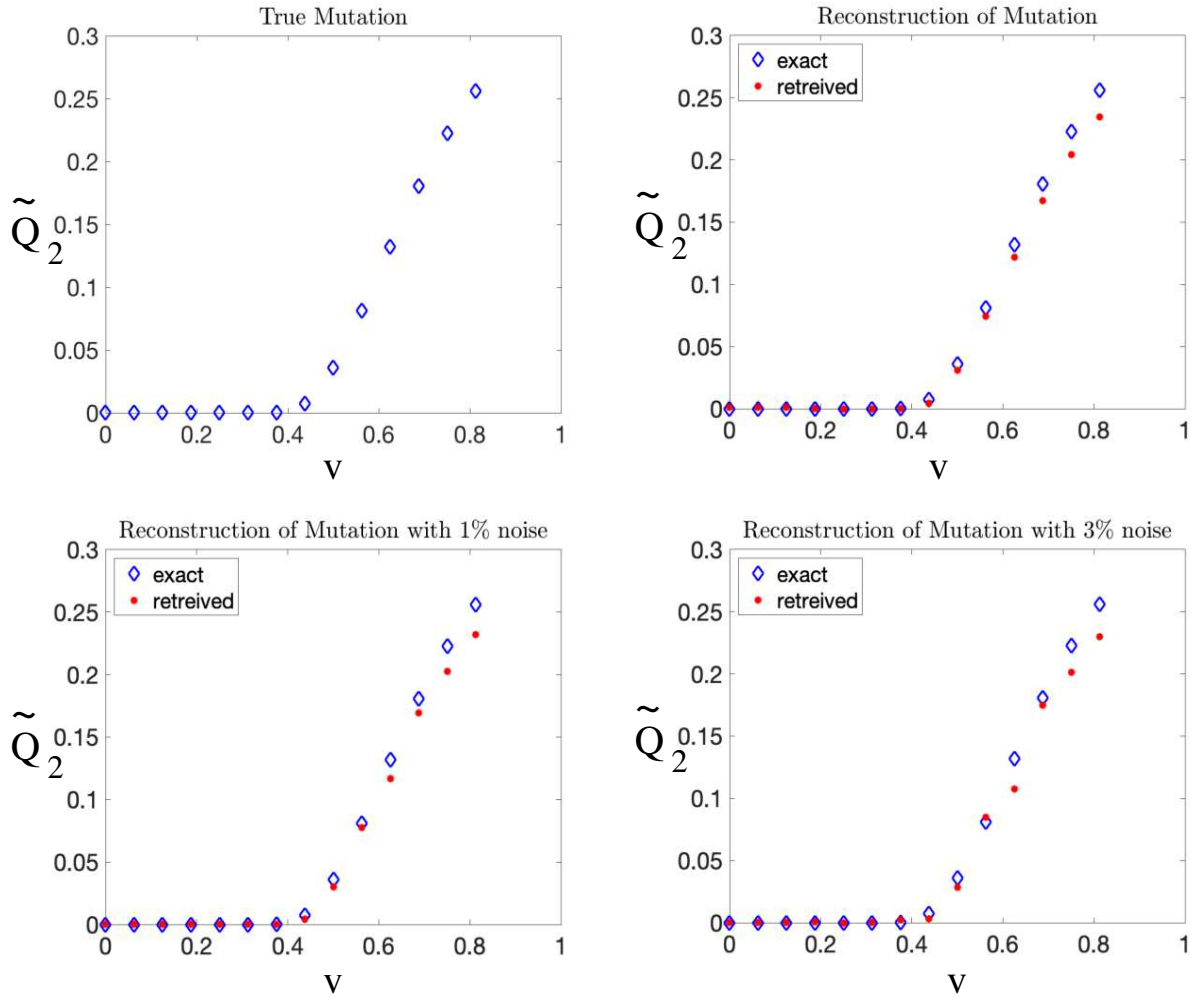

Figure 3. Reconstruction of the unknown mutation law term $\widetilde{Q}_2(v)$ (that appears in mutation law given in model (2.3) for model (2.1). First row: (left) the true mutation law, and (right) the reconstructed mutation law restricted to $\mathcal{A}_v$ with no noise and $\alpha^* = 10^{-12}$. Second row: (left) the reconstructed mutation law obtained for 1% noisy data and $\alpha^* = 10^{-12}$; and (right) the reconstructed mutation law obtained for 3% noisy data and $\alpha^* = 10^{-12}$. For all plots in this figure: 1) the first axis represents the the values for $v \in [\bar{v}^{min}, \bar{v}^{max}]$; 2) second axis represents the values for mutation. The numerical simulations are obtained using the parameters given in Table 1.

Figure 3 shows the reconstruction of unknown nonlinear cancer cell mutation term $\widetilde{Q}_2(v)$ (that is part of cell mutation law introduced in model (2.3)) for model (2.1) in the presence of the measurements given by models (3.15) and (3.18) that are considered here both exact and affected by a level of noise

$\delta \in \{1\%, 3\%\}$. Again, the first row shows from left to right the true mutation law restricted to $\mathcal{A}_v$ where the reconstruction is attempted, and the reconstruction of the mutation law model (2.3) on $\mathcal{A}_v$ with no noise, respectively. The second row of the figure show from left to right the mutation reconstruction on $\mathcal{A}_v$ with 1%, and 3% of noise in the measured data, respectively.

From Figures 2 and 3 we observe that when the measurement data is not affected by noise, we obtain good mutation laws reconstructions in both cases, i.e., (i) mutation depends on primary tumour only, and (ii) mutation depend on the ECM only. However, as expected, as soon as the level of noise in the measurements increases, the reconstruction gradually looses accuracy in both cases (i.e., Case (i) shown in Figure 2 lower panels, and Case (ii) shown in Figure 3 lower panels). This loss in accuracy increases with the increase in the c_1 density (horizontal axis in Figure 2) and in the v density (horizontal axis in Figure 3).

3.4. Reconstruction of nonlinear mutation law in Case (iii)

In this section, we study the inverse problem of identifying the general mutation law in case (iii), namely $Q(c_1, v)$, which is unknown and depends nonlinearly in an autonomous manner on both primary tumour cell density c_1 and on ECM density v.

3.4.1. Inverse problem setup: forward solver computational formulation

Maintaining here the same spatio-temporal discretisation notations introduced in Section 3.1 for the cancer cell populations and ECM, we seek to identify the unknown nonlinear mutation law $Q(c_1, v)$ such that we match the measurements model (3.1). To achieve this, similar to the previous two cases addressed in Sections 3.1−3.3, within an suitable space of functions $\mathfrak{M}^2$ that will be detailed below, we proceed to identify an appropriate a mutation approximating function defined on the 2-dimensional region where the pair (c_1, v) ranges during its spatio-temporal evolution, namely $\overline{m}^{c_1^*, c_2^*, v^*} : [0, K_c] \times [0, K_c] \to [0, \infty)$. The function $\overline{m}^{c_1^*, c_2^*, v^*}$ will be identified such that this will directly determine an acceptable *"mutation law candidate"*, denoted here by $Q^{c_1^*, c_2^*, v^*}(\cdot, \cdot)$, which will enable a solution in *forward model* (2.1) that matches the measurements model (3.1). To select this mutation law candidate $Q^{c_1^*, c_2^*, v^*}(\cdot, \cdot)$, adopting a similar approach to the one in Cases (*i*) and (*ii*), we involve again an appropriately constructed *mutation trial operator* $\mathcal{M}^2(\cdot, \cdot, \cdot) := \{\mathcal{M}^2_{i,j}(\cdot, \cdot, \cdot)\}_{i,j=1\ldots N}$, $\mathcal{M}^2(\cdot, \cdot, \cdot) : \mathbb{R}^{N\times N} \times \mathbb{R}^{N\times N} \times \mathfrak{M}^2 \to \mathbb{R}^{N\times N}$ that will be specified in a moment and will enable us the express the acceptable mutation law candidate as

$$Q^{c_1^*, c_2^*, v^*}(\tilde{c}^{\overline{m}^{c_1^*, c_2^*, v^*}}_{1,i,j}(t), \tilde{v}^{\overline{m}^{c_1^*, c_2^*, v^*}}_{i,j}(t)) := \mathcal{M}^2_{i,j}(\tilde{c}^{\overline{m}^{c_1^*, c_2^*, v^*}}_1(t), \tilde{v}^{\overline{m}^{c_1^*, c_2^*, v^*}}(t), \overline{m}^{c_1^*, c_2^*, v^*}) \tag{3.22}$$

where $\tilde{c}^{\overline{m}^{c_1^*, c_2^*, v^*}}_1(t) := \{\tilde{c}^{\overline{m}^{c_1^*, c_2^*, v^*}}_1(t)\}_{i,j=1\ldots N}$ and $\tilde{v}^{\overline{m}^{c_1^*, c_2^*, v^*}}(t) := \{\tilde{v}^{\overline{m}^{c_1^*, c_2^*, v^*}}(t)\}_{i,j=1\ldots N}$ represents the solution for the density of the primary cell population and of the ECM, respectively, which are obtained for model (2.1) when, instead of the unknown term $Q(\cdot, \cdot)$, in the mutation law we use the trial mutation term $\mathcal{M}^2(\cdot, \cdot, m^{c_1^*, c_2^*, v^*})$

Denoting by $\mathcal{G}^2_M := \{(\eta_l, \zeta_k)\}_{l,k=1\ldots M}$ the equally spaced grid given by the uniform discretisation of the $[0, K_c] \times [0, K_c]$ with step size $\Delta\eta = \Delta\zeta > 0$, the space of functions where we seek to identify the function $\overline{m}^{c_1^*, c_2^*, v^*}$ is an $M \times M$- dimensional space of potential mutation law shape candidates, which is

given by

$$\begin{aligned}\mathfrak{M}^2 :=&\left\{\overline{m} : [0,K_c]\times[0,K_c]\to\mathbb{R}\;\middle|\; \overline{m}|_{E_{l,k}} = \sum_{p,q=0,1} \overline{m}(\eta_{l+p},\zeta_{k+q})\phi_{l+p,k+q},\quad \forall E_{l,k}\in\mathcal{G}_M^{2,tiles}\right\}\\ &\text{with}\quad \mathcal{G}_M^{2,tiles} := \{E_{l,k} := [\eta_l,\eta_{l+1}]\times[\zeta_k,\zeta_{k+1}]\,|\,l,k=1\dots M-1\},\ \textit{and}\\ &\forall\ E_{l,k}\in\mathcal{G}_M^{2,tiles},\ \{\phi_{l+p,k+q}\}_{p,q=0,1}\text{ are the usual bilinear shape functions on } E_{l,k}.\end{aligned} \tag{3.23}$$

Therefore, for any $\overline{m}\in\mathfrak{M}^2$, the *trial proliferation operator* $\mathcal{M}^2$ is given by

$$\begin{aligned}&\mathcal{M}^2_{i,j}(\tilde{c}_1^{\overline{m}}(t),\tilde{v}^{\overline{m}}(t),\overline{m}) := \overline{m}|_{E_{l,k}}(\tilde{c}_1^{\overline{m}}(t),\tilde{v}^{\overline{m}}(t))\\ &\text{with } (l,k)\in\Lambda_{i,j} := \left\{l,\ k\in\{1,\dots,M-1\}\;\middle|\;\exists E_{l',k'}\in\mathcal{G}_M^{tiles}\text{ such that }(\tilde{c}^{\overline{m}}_{1,i,j}(t),\tilde{v}^{\overline{m}}_{i,j}(t))\in E_{l',k'}\right\},\\ &\text{and noting also here that } (l,k) \text{ is independent of its choice within } \Lambda_{i,j}.\end{aligned} \tag{3.24}$$

Here, $\tilde{c}_1^{\overline{m}}(t) := \{\tilde{c}^{\overline{m}}_{1,i,j}(t)\}_{i,j=1\dots N}$, $\tilde{c}_2^{\overline{m}}(t) := \{\tilde{c}^{\overline{m}}_{2,i,j}(t)\}_{i,j=1\dots N}$ and $\tilde{v}^{\overline{m}}(t) := \{\tilde{v}^{\overline{m}}_{i,j}(t)\}_{i,j=1\dots N}$ represent the solutions at the grid points and time $t>0$ for the cancer cells and ECM densities obtained with model (2.1) when this uses $\mathcal{M}^2_{i,j}(\cdot,\cdot,\overline{m})$ as mutation law given in model (3.24). Therefore, in space-discretised form, model (2.1) can be recasted also in this case as

$$\frac{\partial}{\partial t}\begin{bmatrix}\tilde{c}_1^{\overline{m}}\\ \tilde{c}_2^{\overline{m}}\\ \tilde{v}^{\overline{m}}\end{bmatrix} = \begin{bmatrix}\mathcal{H}^1(\tilde{c}_1^{\overline{m}},\tilde{c}_2^{\overline{m}},\tilde{v}^{\overline{m}},\overline{m})\\ \mathcal{H}^2(\tilde{c}_1^{\overline{m}},\tilde{c}_2^{\overline{m}},\tilde{v}^{\overline{m}},\overline{m})\\ \mathcal{H}^3(\tilde{c}_1^{\overline{m}},\tilde{c}_2^{\overline{m}},\tilde{v}^{\overline{m}})\end{bmatrix}, \tag{3.25}$$

Also in in this case (i.e., case (*iii*)) we have that $\mathcal{H}^1(\cdot,\cdot,\cdot,\cdot)$ and $\mathcal{H}^2(\cdot,\cdot,\cdot,\cdot)$ are correspondingly defined through Eqs (3.7) and (3.8) when the *trial mutation form* for the full mutation law (which in Cases (*i*) and (*ii*) was given by operator $\overline{\mathcal{M}}_r$) is given here by $\mathcal{M}^2_{i,j}(\tilde{c}_1^{\overline{m}}(t),\tilde{v}^{\overline{m}}(t),\overline{m})$. Finally, adopting again the same time discretisation as in Section 3.1 and using the Euler method, a time marching step for model (3.25) is given by following operator

$$\overline{K}_{\overline{m}} : \mathbb{R}^{N\times N}\times\mathbb{R}^{N\times N}\times\mathbb{R}^{N\times N}\to\mathbb{R}^{N\times N}\times\mathbb{R}^{N\times N}\times\mathbb{R}^{N\times N}$$

given by

$$\overline{K}_{\overline{m}}\left(\begin{bmatrix}\tilde{c}_1^{m_1,n}\\ \tilde{c}_2^{m_1,n}\\ \tilde{v}^{m_1,n}\end{bmatrix}\right) := \begin{bmatrix}\tilde{c}_1^{m_1,n}\\ \tilde{c}_2^{m_1,n}\\ \tilde{v}^{m_1,n}\end{bmatrix} + \Delta t\begin{bmatrix}\mathcal{H}^1(\tilde{c}_1^{\overline{m},n},\tilde{c}_2^{\overline{m},n},\tilde{v}^{\overline{m},n},\overline{m})\\ \mathcal{H}^2(\tilde{c}_1^{\overline{m},n},\tilde{c}_2^{\overline{m},n},\tilde{v}^{\overline{m},n},\overline{m})\\ \mathcal{H}^3(\tilde{c}_1^{\overline{m},n},\tilde{c}_2^{\overline{m},n},\tilde{v}^{\overline{m},n})\end{bmatrix}, \tag{3.26}$$

where, for any $n\in\{0,\dots,L\}$, we have $\tilde{c}_1^{\overline{m},n} := \tilde{c}_1^{\overline{m}}(t_n)$, $\tilde{c}_2^{\overline{m},n} := \tilde{c}_2^{\overline{m}}(t_n)$, and $\tilde{v}^{\overline{m},n} := \tilde{v}^{\overline{m}}(t_n)$, while

$$\begin{aligned}&\mathcal{H}^1(\tilde{c}_1^{\overline{m},n},\tilde{c}_2^{\overline{m},n},\tilde{v}^{\overline{m},n},\overline{m}) := \mathcal{H}^1(\tilde{c}_1^{\overline{m}}(t_n),c_2^{\overline{m}}(t_n),\tilde{v}^{\overline{m}}(t_n),\overline{m}),\\ &\mathcal{H}^2(\tilde{c}_1^{\overline{m},n},\tilde{c}_2^{\overline{m},n},\tilde{v}^{\overline{m},n},\overline{m}) := \mathcal{H}^2(\tilde{c}_1^{\overline{m}}(t_n),c_2^{\overline{m}}(t_n),\tilde{v}^{\overline{m}}(t_n),\overline{m}),\\ &\mathcal{H}^3(\tilde{c}_1^{\overline{m},n},\tilde{c}_2^{\overline{m},n},\tilde{v}^{\overline{m},n}) := \mathcal{H}^3(\tilde{c}_1^{\overline{m}}(t_n),\tilde{c}_2^{\overline{m}}(t_n),\tilde{v}^{\overline{m}}(t_n)).\end{aligned}$$

Therefore, this allows us formulate *"forward operator"* denoted here by $\overline{K}$ defined by

$$\overline{K} : \mathfrak{M}^2 \to \mathbb{R}^{N\times N} \times \mathbb{R}^{N\times N} \times \mathbb{R}^{N\times N}$$

given by

$$\overline{K}(\overline{m}) := \underbrace{\overline{K}_{\overline{m}} \circ \overline{K}_{\overline{m}} \circ \cdots\cdots \circ \overline{K}_{\overline{m}}}_{L \text{ times}} \left(\begin{bmatrix} \tilde{c}_{1,0} \\ \tilde{c}_{2,0} \\ \tilde{v}_0 \end{bmatrix} \right) \tag{3.27}$$

where $\tilde{c}_{1,0}$, $\tilde{c}_{2,0}$ and $\tilde{v}_0$ are the discretised initial conditions for model (2.1) (as introduced in model (3.11)).

Similar to the previous two inverse problem cases, the *forward operator* $\overline{K}$ will enable us to identify the $\overline{m}^{c_1^*,c_2^*,v^*} \in \mathfrak{M}^2$ such that the solution of the resulting model matches measurements model (3.1). Thus, similar to Section 3.2, we observe again that $\overline{K}$ can be written down as

$$\mathfrak{M}^2 \ni \overline{m} \longmapsto \overline{K}_{\overline{m}} \in \boldsymbol{\ell}^2(\boldsymbol{\ell}^2(\mathcal{E}\times\mathcal{E}\times\mathcal{E}); \boldsymbol{\ell}^2(\mathcal{E}\times\mathcal{E}\times\mathcal{E})) \tag{3.28}$$

Here, $\boldsymbol{\ell}^2(\boldsymbol{\ell}^2(\mathcal{E}\times\mathcal{E}\times\mathcal{E}); \boldsymbol{\ell}^2(\mathcal{E}\times\mathcal{E}\times\mathcal{E}))$ is again the finite-dimensional Bochner space of square integrable vector-value functions defined on $\boldsymbol{\ell}^2(\mathcal{E}\times\mathcal{E}\times\mathcal{E})$ with values in the same space. As in Section 3.2, also here we have that the mappings defined in model (3.28) are continuous and compact. Therefore, we obtain that $\overline{K}$ is also closed sequentially bounded, and as a consequence we have that also in this case we satisfy the inverse problems hypotheses adopted [30], and thus, the convergence of the subsequent nonlinear Tikhonov regularisation is ensured. The Tikhonov functionals $\{\overline{J}_\alpha\}_{\alpha>0}$ will have essentially the same form as in the first two cases, except that the space where these are defined is different (i.e., in this case we have $\mathfrak{M}^2$ of dimension $M \times M$ rather that $\mathfrak{M}^1$ of dimension M that we had in Cases (*i*) and (*ii*)), namely:

$$\overline{J}_\alpha : \mathfrak{M}^2 \to \mathbb{R}, \quad \forall\, \alpha > 0,$$

defined by

$$\overline{J}_\alpha(\overline{m}) := \left\| \overline{K}(\overline{m}) - \begin{bmatrix} \tilde{c}_1^* \\ \tilde{c}_2^* \\ \tilde{v}^* \end{bmatrix} \right\|_2^2 + \alpha \| \overline{m} \|_2^2, \quad \forall\, \overline{m} \in \mathfrak{M}^2. \tag{3.29}$$

3.4.2. Numerical reconstruction of the mutation law $Q(c_1, v)$ in Case (iii)

Computationally, we address here the reconstruction of the general mutation law $Q(c_1, v)$ that appears in Case (*iii*) for model (2.1) in the presence of zero flux boundary conditions and the initial conditions prescribed in Eqs (3.17a)–(3.17c). Furthermore, the inverse problem is addressed in the presence of both exact and noisy measurements of the form detailed in Eqs (3.15) and (3.16), with the exact measurements given by

$$\tilde{c}_{1,exact}^*(x) := \bar{c}_1(x, t_f), \quad \tilde{c}_{2,exact}^*(x) := \bar{c}_2(x, t_f), \quad \textit{and} \quad \tilde{v}_{exact}^*(x) := \bar{v}(x, t_f), \quad \forall x \in \Omega, \tag{3.30}$$

where $\bar{c}_1(x,t)$, $\bar{c}_2(x,t)$ and $\bar{v}(x,t)$ represent the solution densities for primary cell population, mutated cell population and ECM, respectively, which are obtained when model (2.1) uses as mutation law the expression given by model (2.3) with the known term $\widetilde{Q}_2(v)$ specified in model (2.4).

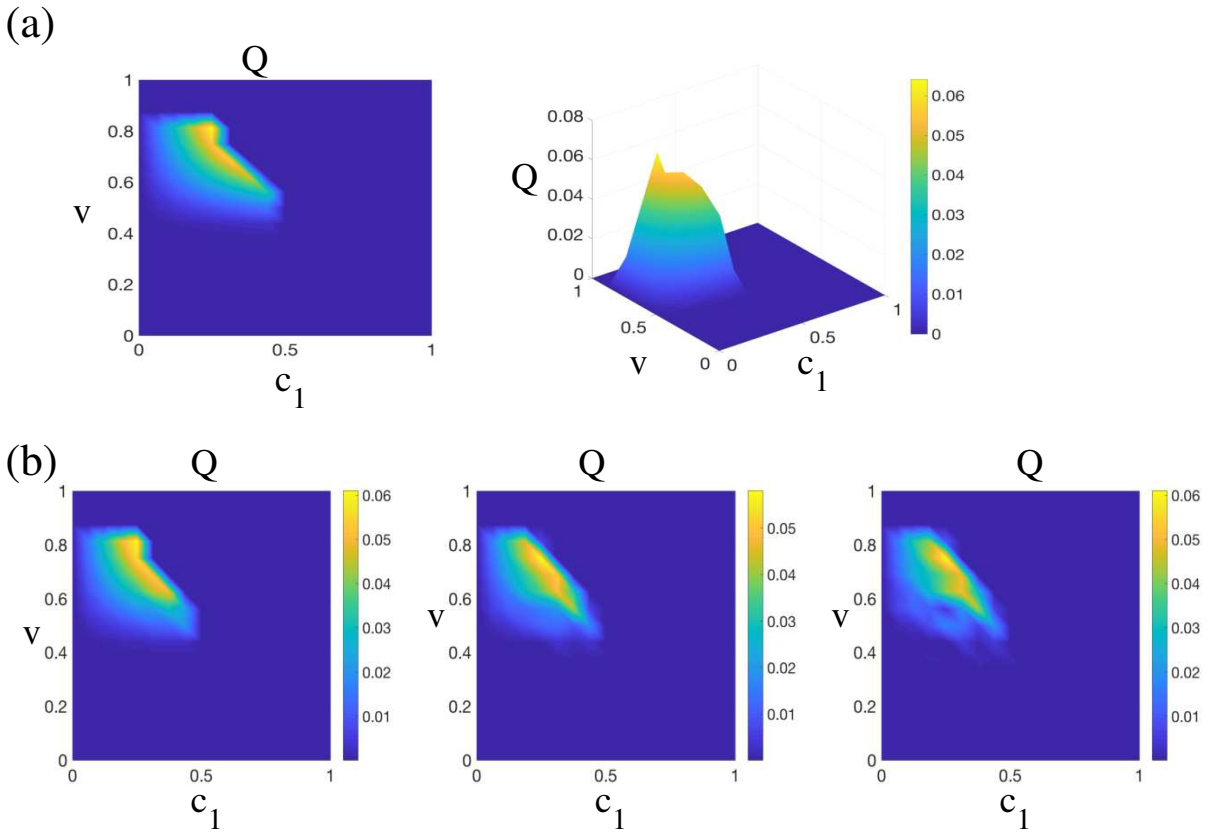

Figure 4. Reconstruction of the general mutation law $Q(c_1, v)$ in Case (*iii*) for model (2.1). Row a) the true mutation law restricted to $\mathcal{A}_{cv}$; Row b) the reconstructed mutation law on $\mathcal{A}_{cv}$ in the presence of exact and noisy data: (left) exact data and $\alpha^* = 10^{-8}$; (centre) 1% noisy data and $\alpha^* = 10^{-5}$; and (right) 3% noisy data and $\alpha^* = 10^{-5}$. For all plots in this figure: 1) first axis shows the values $c_1 \in [\bar{c}_1^{min}, \bar{c}_1^{max}]$; 2) second axis shows the values $v \in [\bar{v}^{min}, \bar{v}^{max}]$; and 3) colour bars represent the magnitude of mutation law or its reconstructions at each $(c_1, v) \in \mathcal{A}_{cv}$. Numerical simulations are obtained using the parameters given in Table 1.

Although the dimensionality is different with respect to Cases (*i*) and (*ii*), the inversion method in this case follows the same steps and numerical minimisation procedure and steps that were described in Section 3.3 (reason for which we skip here details already outlined there). Indeed, for any $\alpha \in \{10^{-i} \,|\, i = 1, \dots 12\}$, the minimisation of the Tikhonov functional J_α starts with an initial guess $\overline{m}_0 = \mathbb{I} \times 10^{-3}$, (where $\mathbb{I}$ represents the $M \times M$ matrix of ones), and leads to the numerical identification of the associated point of minimum m^α that correspond to the smalles mismatch between the associated solution (that is obtained when model (2.1) uses m^α as mutation law) and the measurements . Furthermore, since no data is available beyond the maximal region *maximal accessible regions* $\mathcal{A}_{cv}$ given by

$$\mathcal{A}_{cv} := \mathcal{A}_c \times \mathcal{A}_v = [\bar{c}_1^{min}, \bar{c}_1^{max}] \times [\bar{v}^{min}, \bar{v}^{max}], \text{ with } \mathcal{A}_c \text{ and } A_v \text{ described in Eqs (3.19) and (3.20).}$$

the reconstruction of the mutation law is explored only on $\mathcal{A}_{cv}$. Finally, using again an discrepancy principle-based argument to choose the regularisation parameter $\alpha* \in \{10^{-i} \,|\, i = 1, \dots 12\}$, the reconstructed mutation law will be given by the corresponding $M \times M$ matrix $\overline{m}^{c_1^*, c_2^*, v^*} := \overline{m}^{\alpha*}$, which

in turn will determine the precise shape of the reconstructed mutation law $\mathcal{Q}^{c_1^*,c_2^*,v^*}(\tilde{c}_{1,i,j}^{\overline{m}\,c_1^*,c_2^*,v^*}(t), \tilde{v}_{i,j}^{\overline{m}\,c_1^*,c_2^*,v^*}(t))$ that is defined as per model (3.22). Figure 4 shows the reconstruction of the most general cancer cell mutation law $\mathcal{Q}(c_1, v)$, starting from the measurements given by models (3.15) and (3.18) for exact data as well as data impacted by noise $\delta \in \{1\%, 3\%\}$. For comparison, the first row of this figure shows the true mutation law restricted at the maximal accessible region $\mathcal{A}_{cv}$ where the reconstruction is being attempted. The second row of the figure show from left to right the reconstruction of the mutation law on $\mathcal{A}_{cv}$ with no noise (left), with 1% noise (center), and 3% noise (right) in the measured data. From this figure, we observe that we obtain a good reconstruction of the mutation law when the measurement data are not affected by noise. However, as expected, as soon as the level of noise increases in the measurements, the reconstruction gradually looses accuracy.

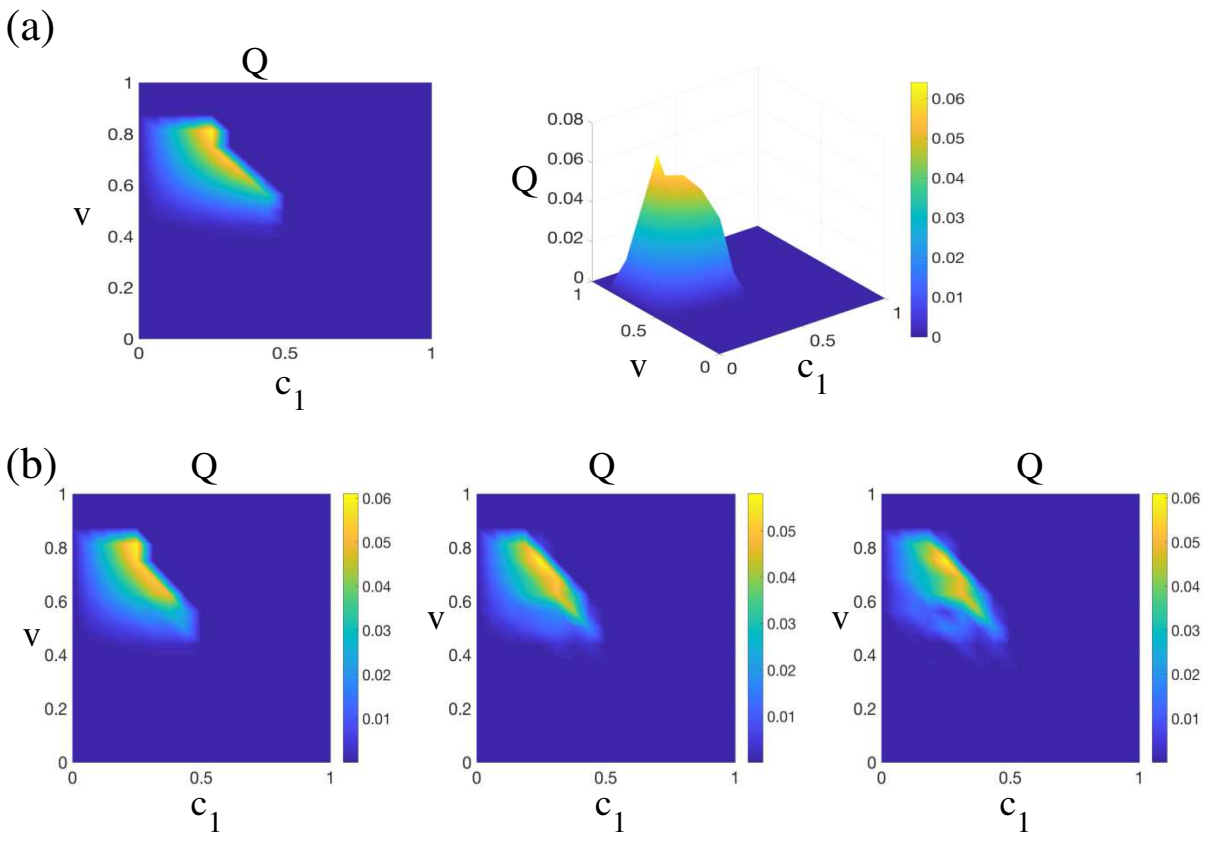

Figure 5. Reconstruction of general mutation law $\mathcal{Q}(c_1, v)$ in Case (*iii*) for model (2.1), obtained when using logistic cell proliferation for c_1 and Gompertz cell proliferation for c_2. Row a): the true mutation law restricted to $\mathcal{A}_{cv}$; Row b) the reconstructed mutation law on $\mathcal{A}_{cv}$ in the presence of exact and noisy data: (left) exact data and $\alpha^* = 10^{-4}$; (centre) 1% noisy data and $\alpha^* = 10^{-4}$; and (right) 3% noisy data and $\alpha^* = 10^{-4}$. For all plots in this figure: 1) first axis shows the values $c_1 \in [\bar{c}_1^{min}, \bar{c}_1^{max}]$; 2) second axis shows the values $v \in [\bar{v}^{min}, \bar{v}^{max}]$; and 3) the colour bars represent the magnitude of mutation law or its reconstructions at each $(c_1, v) \in \mathcal{A}_{cv}$. The numerical simulations are obtained with the parameters given in Table 1.

3.4.3. Reconstruction of unknown mutation law $Q(c_1, v)$ in Case (*iii*) for a different cell proliferation rule for c_2

Throughout the previous sections we assumed that both cancer cell populations proliferate logistically: $\mu_c c_{1,2}\left(1\frac{c_1+c_2+v}{K_c}\right)$. However, the sigmoid shape of tumour growth that is given by the logistic term can be obtained also with other proliferation rules, such as the Gompertz rule [33–35]: $\mu_c c_{1,2}\log\left(\frac{K_c}{c_1+c_2+v}\right)$. This raises the question as to what happens when different cancer cell populations use different proliferation laws.

In this subsection we reconstruct the general mutation law $Q(c_1, v, t)$ in case (*iii*) when we assume that the primary c_1 population proliferation proliferates logistically, while the second (mutated) cell population c_2 proliferates according to Gompertz law. In Figure 5 we present the numerical reconstruction results. We observe that the results are similar to those in Figure 4; this could be explained by the fact that the mutation starts in the primary tumour which proliferate logistically.

4. A tumour invasion model with nonlocal dynamics

Since various mathematical studies have assumed nonlocal cell-cell and cell-ECM interactions to explain the invasion of cancer cells [11, 16, 22, 23], in the following we generalise the model (2.1) by replacing the local haptotaxis towards local ECM gradients with a nonlocal haptotaxis flux generated by these cell-cell and cell-ECM interactions.

As before, we consider a primary cancer cell subpopulation $c_1(x, t)$ and a secondary mutated cancer cell subpopulation $c_2(x, t)$. These cancer cell populations interact with each other as well as with the extracellular matrix (ECM), $v(x, t)$, which they degrade and remodel. For compact notation, we consider the combined vector of primary cancer cells c_1, mutated cancer cells c_2 and extracellular matrix v defined as

$$\mathbf{u}(x,t) := [c_1(x,t), c_2(x,t), v(x,t)]^T .$$

We use this vector to describe in a compact manner the flux term generated by the nonlocal cell-cell and cell-ECM interactions ($\mathcal{A}_{1,2}(x, t, \mathbf{u}\cdot, t)$). Therefore the coupled tumour dynamics in this case is given by:

$$\frac{\partial c_1}{\partial t} = \underbrace{D_1\Delta c_1}_{\text{random motility}} - \underbrace{\nabla\cdot\left[c_1\mathcal{A}_1\left(x,t,\mathbf{u}\left(\cdot,t\right)\right)\right]}_{\text{adhesion}} + \underbrace{\mu_c c_1\left(1-\frac{c_1+c_2+v}{K_c}\right)}_{\text{logistic proliferation}} - \underbrace{\omega(t)}_{\substack{\text{mutation}\\ \text{switch}}}\underbrace{Q(\cdot,\cdot)}_{\substack{\text{unknown}\\ \text{mutation}}}, \tag{4.1a}$$

$$\frac{\partial c_2}{\partial t} = \underbrace{D_2\Delta c_2}_{\text{random motility}} - \underbrace{\nabla\cdot\left[c_2\mathcal{A}_2\left(x,t,\mathbf{u}\left(\cdot,t\right)\right)\right]}_{\text{adhesion}} + \underbrace{\mu_c c_2\left(1-\frac{c_1+c_2+v}{K_c}\right)}_{\text{logistic proliferation}} + \underbrace{\omega(t)}_{\substack{\text{mutation}\\ \text{switch}}}\underbrace{Q(\cdot,\cdot)}_{\substack{\text{unknown}\\ \text{mutation}}}, \tag{4.1b}$$

$$\frac{\partial v}{\partial t} = \underbrace{-\rho(c_1+c_2)v}_{\text{degradation}} + \underbrace{\mu_v\left(K_c-c_1-c_2-v\right)^+}_{\text{ECM remodelling}} . \tag{4.1c}$$

As before, $D_{1,2}$ are the diffusion coefficients, μ_c is the cancer cell proliferation rate, ρ is the ECM degradation rate, and μ_v is the ECM remodelling rate. The flux term $\mathcal{A}_p\left(x,t,\mathbf{u}\left(\cdot,t\right)\right)$, $p = 1, 2$, has been proposed in [11, 23] to describe the directed movement of cells due to cell-cell and cell-matrix

adhesion:

$$\mathcal{A}_p\left(x,t,\mathbf{u}\left(\cdot,t\right)\right) := \frac{1}{R}\int\limits_{\mathbf{B}((0,0),R)} n\left(y\right)\cdot\mathcal{K}\left(\|y\|_2\right)\cdot g_p\left(\mathbf{u}\left(x+y,t\right),t\right)\chi_\Omega\left(x+y\right)\,dy. \tag{4.2}$$

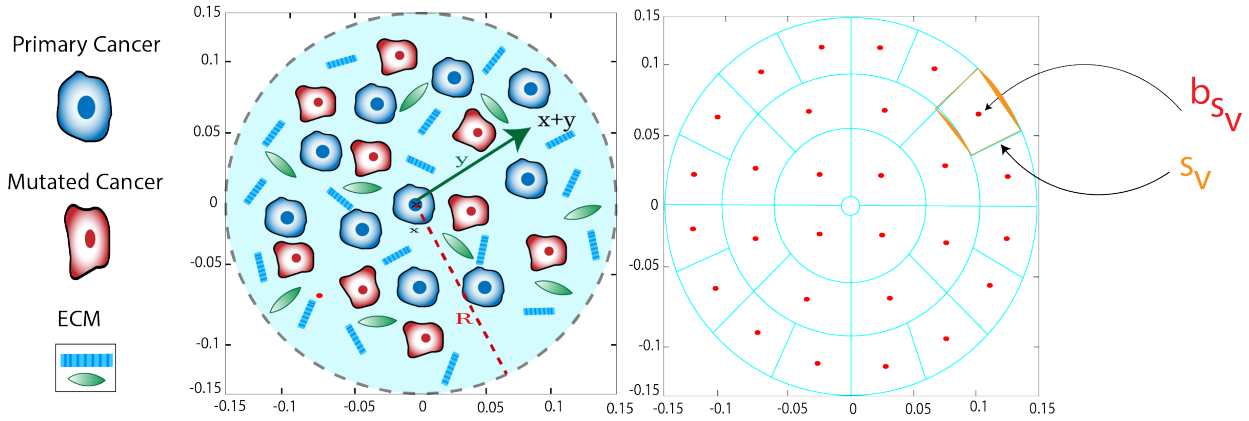

Figure 6. Schematic shows the sensing region $B(0,R)$ and radial direction for cell-cell and cell-matrix adhesion interactions. The left figure describes the process of the cancer cells adhesion, the ball $B(x,R)$ centred at x and of radius R, the point $x+y$ with the direction vector $n(y)$ in green. The right figure shows the decomposition of the region using annulus sectors S_v with barycentres b_{S_v}, highlighted with red dots, which are defined with full details in the Appendix Section A where we describe the numerical approach for the nonlocal forward model (4.1).

By using these notations the approximation for $D_p\Delta c_p - \nabla\cdot[c_p\mathcal{A}_p\left(x,t,\mathbf{u}\left(\cdot,t\right)\right)]$ in (4.1) is as follows:

It is assumed that the interactions between a cell and its neighbouring cells as well as the components of the ECM occur inside a sensing region $\mathbf{B}\left((0,0),R\right)\subset\mathbb{R}$, where $R>0$ is the *sensing radius*, as illustrated in Figure 6. In the above equation, $\chi_\Omega\left(\cdot\right)$ represents the characteristic function of Ω. Further, $n\left(y\right)$ is the unit radial vector given by

$$n\left(y\right) := \begin{cases} y/\|y\|_2 & \text{if} \quad y\in\mathbf{B}\left(0,R\right)\setminus\{(0,0)\},\\ (0,0) & \text{otherwise.}\end{cases} \tag{4.3}$$

Furthermore, $g_p\left(\mathbf{u}(x+y,t),t\right)$ represents the $p-th$ component of the adhesion function $g\left(\mathbf{u}\left(x+y,t\right),t\right)$ that is given by

$$\begin{gathered} g\left(\mathbf{u},t\right) = \left[\mathbf{S}_{cc}\mathbf{c}+\mathbf{S}_{cv}v\right]\cdot\left(K_c-c_1-c_2-v\right)^+,\\ \text{with}\quad \mathbf{S}_{cc}=\begin{pmatrix}S_{c_1c_1} & S_{c_1c_2}\\ S_{c_2c_2} & S_{c_2c_2}\end{pmatrix}\quad\text{and}\quad \mathbf{S}_{cv}=\begin{pmatrix}S_{c_1v} & 0\\ 0 & S_{c_2v}\end{pmatrix}. \end{gathered} \tag{4.4}$$

Here $\{S_{c_ic_j}\}_{i=1,2}$ are the non-negative cell-cell adhesion strengths of the adhesion bonds established between the primary and mutated cancer cell populations, while $\{S_{c_iv}\}_{i=1,2}$ are the non-negative

cell-matrix adhesion stands for the strengths for the adhesive interactions between each of the two cancer subpopulation and the ECM. Finally, no overcrowding of the cell population and ECM over the tumour domain is ensured here through the term $(K_c - c_1 - c_2 - v)^+ := \max(K_c - c_1 - c_2 - v, 0)$.

The nonlocal dynamics model (4.1) is assumed to take place in the presence of the same initial and boundary condition as those assumed in for the local model and given by models (2.5) and (2.6) in Section 2. The numerical approximation of the forward cancer invasion model (4.1) is presented in Appendix Section A. This includes the discretisation of the main spatial operators as well as the off grid numerical approach for the nonlocal adhesion flux terms $\mathcal{A}_p(x, t, \mathbf{u}(\cdot, t))$ whose schematic is given by the right side of Figure 6. Finally, the general mutation law $Q(c_1, v)$ is assumed to be unknown and its identification will be our main focus in this section.

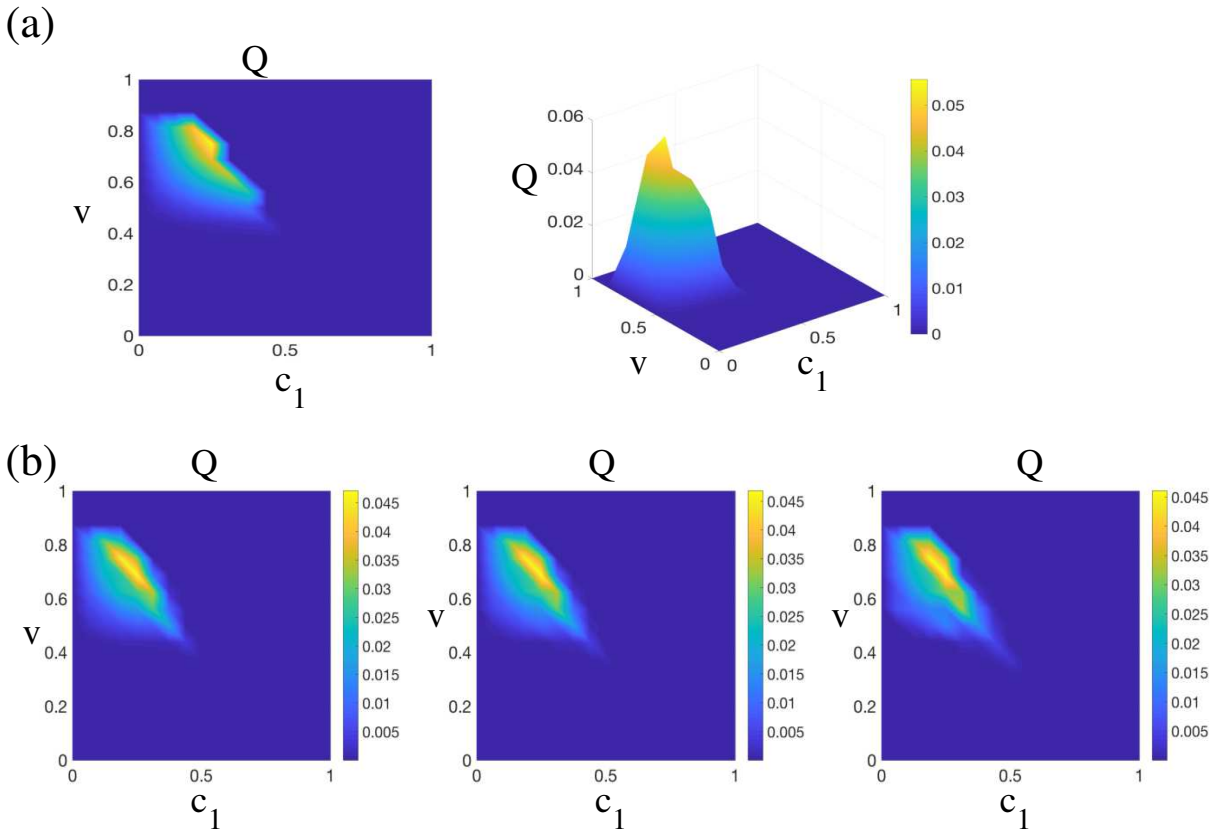

Figure 7. Reconstruction of the general mutation law $Q(c_1, v)$ for the nonlocal model (4.1). Row a) the true mutation law restricted to $\mathcal{A}_{cv}$. Row b) the reconstructed mutation law on $\mathcal{A}_{cv}$ in the presence of exact and noisy data: (left) exact data and $\alpha^* = 10^{-4}$; (centre) 1% noisy data and $\alpha^* = 10^{-4}$; and (right) 3% noisy data and $\alpha^* = 10^{-4}$. For all plots in this figure: 1) first axis shows the values $c_1 \in [\bar{c}_1^{min}, \bar{c}_1^{max}]$; 2) second axis shows the values $v \in [\bar{v}^{min}, \bar{v}^{max}]$; and 3) the colour bars represent the magnitude of mutation law or its reconstructions at each $(c_1, v) \in \mathcal{A}_{cv}$. The numerical simulations are obtained using the parameters given in Table 1.

4.1. Reconstruction of the mutation law in the nonlocal cancer invasion model

Building on the inverse problems approach developed for the local cancer invasion model in Section 3.4 and applying this to the nonlocal model (4.1), we proceed now to address the reconstruction of the unknown mutation law $Q(c_1, v)$ within the nonlocal cancer invasion model (4.1) from exact and noisy measured data that are considered here again to be of the form prescribed in models (3.15) and (3.16). Here, the exact measurement data that are given by

$$\tilde{c}^*_{1,exact}(x) := \bar{c}_1(x, t_f), \quad \tilde{c}^*_{2,exact}(x) := \bar{c}_2(x, t_f), \quad \textit{and} \quad \tilde{v}^*_{exact}(x) := \bar{v}(x, t_f), \quad \forall x \in \Omega, \tag{4.5}$$

where $\bar{c}_p(x, t)$ $p = 1, 2$ and $\bar{v}(x, t)$ are the corresponding solution densities of the primary and mutated cancer cell populations as well as that of ECM that are obtained for model (4.1) obtained when this uses as mutation law given by model (2.3) with the known term $\widetilde{Q}_2(v)$ detailed in model (2.4). For the numerical reconstruction we consider here the same initial and boundary conditions as in Section 3.3. Figure 7 shows the reconstruction of the cancer cell mutation law $Q(c_1, v)$ for model (4.1) in the presence of the measurements given by model (4.5) that are considered here both exact and affected by a level of noise $\delta \in \{1\%, 3\%\}$. As for the previous figures, the first row shows the true mutation law restricted at the maximal accessible region $\mathcal{A}_{cv}$ where the reconstruction is being attempted. The second row of the figure shows, from left to right, the reconstruction of the mutation law on $\mathcal{A}_{cv}$ with no noise, with 1%, and with 3% of noise in the measured data, respectively. From this figure, we observe that we obtain a good mutation law reconstruction when the measurement data are not affected by noise. However, as expected, as soon as the level of noise increases in the measurements, the reconstruction gradually looses accuracy.

4.2. Reconstruction of the mutation law for a different cell proliferation rules for c_2 population

As in Section 3.4.3, here we investigate the reconstruction of the general mutation law $Q(c_1, v)$ when cancer cell growth is described by different rules: logistic proliferation for the c_1 cells, and Gompertz proliferation for the second (mutated) cancer cell population c_2. Again, we do not see any significant differences between the reconstruction of mutation law in this case (see Figure 8) and in the previous case where both cell populations grow logistically (see Figure 7). We believe that is because in both cases, the population that starts mutating (i.e., the c_1 population) grows logistically.

5. Conclusions

In this work we explored a new inverse problem that addresses the reconstruction of the cancer cells mutation law in two heterotypic cancer invasion models: a model with a local cancer cell haptotaxis flux towards ECM (see model (2.1)), and a model with a nonlocal haptotaxis flux generated by cell-cell and cell-ECM adhesion forces (see model (4.1)). For the reconstruction of this mutation law through an inverse problem Tikhonov regularisation-based approach, we used a numerically-generated spatial tumour snapshot data assumed to be acquired at a later stage in the tumour evolution (in practice, the data can be provided through a medical imaging scan).

This inverse problem approach was implemented computationally via a mixed finite differences - finite element numerical scheme. Specifically, on one hand, we used a Crank-Nicholson-type finite difference scheme for the discretisation of the *forward models* that arises in each of the two tumour

invasion dynamics considered here (i.e., local cancer cell invasion, and nonlocal cancer cell invasion). On the other hand, we developed a finite element approach involving a bilinear shape functions on a square mesh for the discretisation of mutation law candidates recruited from a proposed space of functions $\mathcal{S}$, as well as for their evaluation on a maximal accessible regions where the mutation law reconstruction was performed. Finally, these two parts were appropriately assembled in an optimisation solver that sought to reconstruct the cancer cell mutation law by minimising over the space $\mathcal{S}$ the emerging Tikhonov functionals that were formulated in each of the two cases considered.

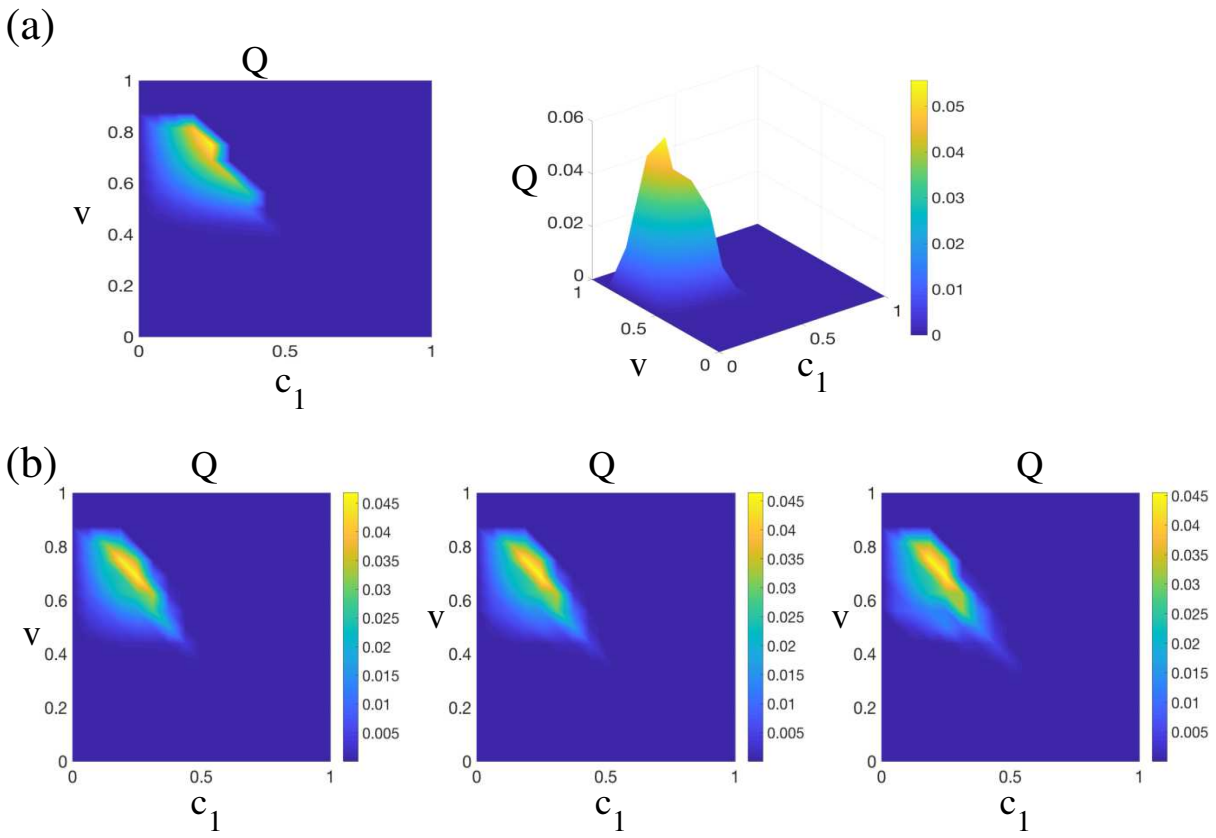

Figure 8. Reconstruction of the general mutation law $Q(c_1, v)$ for the nonlocal model (4.1) with mixed proliferation rules for c_1 and c_2 cells. Row a) the true mutation law restricted to $\mathcal{A}_{cv}$. Row b) the reconstructed mutation law on $\mathcal{A}_{cv}$ in the presence of exact and noisy data: (left) exact data and $\alpha^* = 10^{-4}$; (centre) 1% noisy data and $\alpha^* = 10^{-4}$; and (right) 3% noisy data and $\alpha^* = 10^{-4}$. For all plots in this figure we have that: 1) the first axis represents the values for $c_1 \in [\bar{c}_1^{min}, \bar{c}_1^{max}]$; 2) second axis represents the values for $v \in [\bar{v}^{min}, \bar{v}^{max}]$; and 3) the colour bars represent the magnitude of mutation law or its reconstructions at each $(c_1, v) \in \mathcal{A}_{cv}$. The parameters used for the numerical simulations are given in Table 1.

Finally, this inversion approach was explored and tested on the reconstruction of different cancer cell mutation laws used in local and nonlocal cancer modelling: (i) reconstructing the cancer cell mutation assuming a linear dependence only on the c_1 cell population; (ii) reconstructing the cancer cell mutation assuming a linear dependence on the c_1 population and a nonlinear dependence on the

ECM density, and (iii) reconstructing a very general mutation law, while assuming no prior knowledge about the mutation. To explore the robustness of the proposed inverse problem approach we used *"in silico"* constructed data that mathematically could be expressed in the general form given by models (3.15) and (3.16), but that were specified individually for each of the cases considered for the local and non-local models (2.1) and (4.1), respectively. Indeed, while the noise was always added the same way (as per model (3.16)), the exact measurements considered here were maps given by solutions at the final time for the densities of the cancer cell populations (i.e, both primary and mutated populations) and the ECM, namely $(\bar{c}_1(\cdot, t_f), \bar{c}_2(\cdot, t_f), \bar{v}(\cdot, t_f))$, which were obtained by equipping the local and nonlocal forward models with a selection of known mutation laws (given either by model (2.2) or by models (2.3) and (2.4)) considered as appropriate for each of the cases explored in the context of both local and nonlocal models for tumour dynamics. We note that the results in Figures 4, 5, 7, and 8 look correspondingly similar because we reconstruct the same mutation law Q(c,v), while for the second (mutated) cancer cell population, c_2, we use logistic proliferation for Figures 4 and 7 and Gompertz proliferation law for Figures 5 and 8. While there are small numerical differences, these are hard to be distinguished among the figures without zooming in. Furthermore, this similarity of the results is not significantly affected when switching between the Logistic and Gompertz proliferation laws for the second cancer cells population. An explanation for this similarity is that in addition to that fact that the mutation starts in the primary tumour which proliferate logistically, the logistic and Gompertz laws considered for c_2 are part of the same class of size growth laws [33–35], with similar profiles on the parts of their domain where these are evaluated by the cancer models (2.1) and (4.1). All numerical reconstruction results for the local and nonlocal models showed that for exact measurements we obtained a good reconstruction of the mutation law, while for increasingly noisy measurements the reconstruction gradually deteriorates. Finally, these *in silico* reconstruction tests recommend the inversion framework developed in this work also for the context of real measurements data that come in the form of imaging data (such as MRI) with levels of noise that may not be *a priori* known. Nevertheless, the accuracy of the reconstruction in the real data case is expected to deteriorate for significant levels of noise.

Acknowledgements

The first author would like to thank Saudi Arabian Cultural Bureau in the United Kingdom (UKSACB) on behalf of Taif University and the University College of Al-Khurmah in Saudi Arabia for supporting and sponsoring his PhD studies at the University of Dundee.

Conflict of interest

The authors declare no conflict of interest.

References

1. O. Ikediobi, H. Davies, G. Bignell, S. Edkins, C. Stevens, S. O'Meara, et al., Mutation analysis of 24 known cancer genes in the NCI-60 cell line set, *Mol. Cancer Ther.*, **5** (2006), 2602–2612. https://doi.org/10.1158/1535-7163.MCT-06-0433

2. M. Pickup, J. Mouw, V. Weaver, The extracellular matrix modulates the hallmarks of cancer, *EMBO Rep.*, **15** (2014), 1243–1253. https://doi.org/10.15252/embr.201439246
3. A. López-Carrasco, S. Martín-Vañó, R. Burgos-Panadero, E. Monferrer, A. P. Berbegall, B. Fernández-Blanco, et al., Impact of extracellular matrix stiffness on genomic heterogeneity in mycn-amplified neuroblastoma cell line, *J. Exp. Clin. Cancer Res.*, **39** (2020), 226. https://doi.org/10.1186/s13046-020-01729-1
4. A. L. Jackson, L. A. Loeb, The mutation rate and cancer, *Genetics*, **148** (1998), 1483–1490. https://doi.org/10.1093/genetics/148.4.1483
5. N. Novikov, S. Zolotaryova, A. Gautreau, E. Denisov, Mutational drivers of cancer cell migration and invasion, *Br. J. Cancer*, **124** (2021), 102–114. https://doi.org/10.1038/s41416-020-01149-0
6. A. R. A. Anderson, M. A. J. Chaplain, E. L. Newman, R. J. C. Steele, A. M. Thompson, Mathematical modelling of tumour invasion and metastasis, *J. Theor. Med.*, **2** (2000), 490902. https://doi.org/10.1080/10273660008833042
7. A. R. A. Anderson, A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion, *Math. Med. Biol.*, **22** (2005), 163–186. https://doi.org/10.1093/imammb/dqi005
8. A. R. Anderson, M. Chaplain, Continuous and discrete mathematical models of tumor-induced angiogenesis, *Bull. Math. Biol.*, **60** (1998), 857–899. https://doi.org/10.1006/bulm.1998.0042
9. R. A. Gatenby, E. T. Gawlinski, A reaction-diffusion model of cancer invasion, *Cancer Res.*, **56** (1996), 5745–5753. Available from: https://cancerres.aacrjournals.org/content/56/24/5745.full-text.pdf.
10. M. Chaplain, G. Lolas, Mathematical modelling of cancer cell invasion of tissue: the role of the urokinase plasminogen activation system, *Math. Models Methods Appl. Sci.*, **15** (2005), 1685–1734. https://doi.org/10.1142/S0218202505000947
11. P. Domschke, D. Trucu, A. Gerisch, M. Chaplain, Mathematical modelling of cancer invasion: Implications of cell adhesion variability for tumour infiltrative growth patterns, *J. Theor. Biol.*, **361** (2014), 41–60. https://doi.org/10.1016/j.jtbi.2014.07.010
12. I. Ramis-Conde, D. Drasdo, A. R. Anderson, M. A. Chaplain, Modeling the influence of the e-cadherin-beta-catenin pathway in cancer cell invasion: a multiscale approach, *Biophys. J.*, **95** (2008), 155–165. https://doi.org/10.1529/biophysj.107.114678
13. A. Marciniak-Czochra, M. Ptashnyk, Derivation of a macroscopic receptor-based model using homogenization techniques, *SIAM J. Math. Anal.*, **40** (2008), 215–237. https://doi.org/10.1137/050645269
14. P. Macklin, S. McDougall, A. R. A. Anderson, M. A. J. Chaplain, V. Cristini, J. Lowengrub, Multiscale modelling and nonlinear simulation of vascular tumour growth, *J. Math. Biol.*, **58** (2009), 765–798. https://doi.org/10.1007/s00285-008-0216-9
15. T. S. Deisboeck, Z. Wang, P. Macklin, V. Cristini, Multiscale cancer modeling, *Ann. Rev. Biomed. Eng.*, **13** (2011), 127–155. https://doi.org/10.1146/annurev-bioeng-071910-124729
16. D. Trucu, P. Lin, M. A. J. Chaplain, Y. Wang, A multiscale moving boundary model arising in cancer invasion, *Multiscale Model. Simul. SIAM Int. J.*, **11** (2013), 309–335. https://doi.org/10.1137/110839011

17. T. Colin, A. Iollo, J.-B. Lagaert, O. Saut, An inverse problem for the recovery of the vascularisation of a tumour, *J. Inverse Ill-Posed Probl.*, **22** (2014), 759–786. https://doi.org/10.1515/jip-2013-0009

18. A. Gholami, A. Mang, G. Biros, An inverse problem formulation for parameter estimation of a reaction-diffusion model of low grade gliomas, *J. Math. Biol.*, **72** (2016), 409–433. https://doi.org/10.1007/s00285-015-0888-x

19. C. Hogea, C. Davatzikos, G. Biros, An image-driven parameter estimation problem for a reaction-diffusion glioma growth model with mass effects, *J. Math. Biol.*, **56** (2008), 793–825. https://doi.org/10.1007/s00285-007-0139-x

20. R. Jaroudi, G. Baravdish, B. Johansson, F. Aström, Numerical reconstruction of brain tumours, *Inverse Probl. Sci. Eng.*, **27** (2019), 278–298. https://doi.org/10.1080/17415977.2018.1456537

21. S. Subramanian, K. Scheufele, M. Mehl, G. Biros, Where did the tumour start? An inverse solver with sparse localisation for tumour growth models, *Inverse Probl.*, **36** (2020), 045006. https://doi.org/10.1088/1361-6420/ab649c

22. N. J. Armstrong, K. J. Painter, J. A. Sherratt, A continuum approach to modelling cell-cell adhesion, *J. Theor. Biol.*, **243** (2006), 98–113. https://doi.org/10.1016/j.jtbi.2006.05.030

23. A. Gerisch, M. Chaplain, Mathematical modelling of cancer cell invasion of tissue: Local and non-local models and the effect of adhesion, *J. Theor. Biol.*, **250** (2008), 684–704. https://doi.org/10.1016/j.jtbi.2007.10.026

24. V. Bhandari, C. H. Li, R. G. Bristow, P. C. Boutros, P. Consortium, Divergent mutational processes distinguish hypoxic and normoxic tumours, *Nat. Commun.*, **11** (2020), 737. https://doi.org/10.1038/s41467-019-14052-x

25. F. G. Sonugür, H. Akbulut, The role of tumor microenvironment in genomic instability of malignant tumors, *Front. Genet.*, **10** (2019), 1063. https://doi.org/10.3389/fgene.2019.01063

26. R. Shuttleworth, D. Trucu, Cell-scale degradation of peritumoural extracellular matrix fibre network and its role within tissue-scale cancer invasion, *Bull. Math. Biol.*, **82** (2020), 65. https://doi.org/10.1007/s11538-020-00732-z

27. R. Shuttleworth, D. Trucu, Multiscale dynamics of a heterotypic cancer cell population within a fibrous extracellular matrix, *J. Theor. Biol.*, **486** (2020), 110040. https://doi.org/10.1016/j.jtbi.2019.110040

28. K. Yosida, *Functional Analysis*, Springer-Verlag, 1980.

29. R. L. Schilling, *Measures, Integrals and Martingales*, Cambridge University Press, 2005. https://doi.org/10.1017/CBO9780511810886

30. H. W. Engl, K. Kunisch, A. Neubauer, Convergence rates for tikhonov regularisation of non-linear ill-posed problems, *Inverse Probl.*, **5** (1989), 523–540. https://doi.org/10.1088/0266-5611/5/4/007

31. T. J. R. Hughes, *The Finite Element Method: Linear Static and Dynamics Finite Element Analysis*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1987.

32. V. A. Morozov, *Methods for Solving Incorrectly Posed Problems*, Springer-Verlag, New York, 1984. https://doi.org/10.1007/978-1-4612-5280-1

33. C. Guiot, P. Degiorgis, P. Delsanto, P. Gabriele, T. Diesboeck, Does tumour growth follow a "universal law"? *J. Theor. Biol.*, **225** (2003), 147–151. https://doi.org/10.1016/S0022-5193(03)00221-2

34. A. Laird, Dynamics of tumour growth, *Br. J. Cancer*, **18** (1964), 490–502. https://doi.org/10.1038/bjc.1964.55

35. K. M. C. Tjorve, E. Tjorve, The use of Gompertz models in growth analyses, and new Gompertz-model approach: An addition to the Unified-Richards family, *PLoS One*, **12** (2017), 1–17. https://doi.org/10.1371/journal.pone.0178691

36. R. J. LeVeque, *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems*, SIAM, 2007.

37. L. Peng, D. Trucu, P. Lin, A. Thompson, M. A. Chaplain, A multiscale mathematical model of tumour invasive growth, *Bull. Math. Biol.*, **79** (2017), 389–429. https://doi.org/10.1007/s11538-016-0237-2

38. R. Shuttleworth, D. Trucu, Two-scale moving boundary dynamics of cancer invasion: Heterotypic cell populations' evolution in heterogeneous ecm, in *Cell Movement Modelling and Applications* (eds. M. Stolarska and N. Tarfulea), Birkhauser, Springer Nature Switzerland AG, (2018), 1–26.

39. R. Shuttleworth, D. Trucu, Multiscale modelling of fibres dynamics and cell adhesion within moving boundary cancer invasion, *Bull. Math. Biol.*, **81** (2019), 2176–2219. https://doi.org/10.1007/s11538-019-00598-w

40. V. Andasari, A. Gerisch, G. Lolas, A. P. South, M. A. Chaplain, Mathematical modeling of cancer cell invasion of tissue: biological insight from mathematical analysis and computational simulation, *J. Math. Biol.*, **63** (2011), 141–171. https://doi.org/10.1007/s00285-010-0369-1

Appendix

A. Numerical approach for the forward nonlocal cancer invasion model

In this section, we briefly discuss the numerical methods used to solve the forward models (4.1a)–(4.1c). To discretise the system in space we use the method of lines approach. For the time-evolution of the system, we use a predictor-corrector scheme introduced in [16], using the Euler method as the predictor and a trapezoidal-type rule as the corrector. In the reaction-diffusion models (4.1a) and (4.1b), the terms $D_p\Delta c_p - \nabla\cdot[c_p\mathcal{A}_p(x,t,\mathbf{u}(\cdot,t))]$, $p=1,2$, will be approximated through a second-order mid-point rule [36] as detailed below. In brief, $\forall p\in\{1,2\}$, $n=\overline{0,L}$, $i,j=\overline{1,N}$, these midpoint approximations are given by:

$$\begin{cases} c^n_{p,i,j+\frac{1}{2}} := \frac{c^n_{p,i,j}+c^n_{p,i,j+1}}{2}, \\ c^n_{p,i,j-\frac{1}{2}} := \frac{c^n_{p,i,j}+c^n_{p,i,j-1}}{2}, \\ c^n_{p,i+\frac{1}{2},j} := \frac{c^n_{p,i,j}+c^n_{p,i+1,j}}{2}, \\ c^n_{p,i-\frac{1}{2},j} := \frac{c^n_{p,i,j}+c^n_{p,i-1,j}}{2}. \end{cases} \quad \text{and} \quad \begin{cases} \mathcal{A}^n_{p,i,j+\frac{1}{2}} := \frac{\mathcal{A}^n_{p,i,j}+\mathcal{A}^n_{p,i,j+1}}{2}, \\ \mathcal{A}^n_{p,i,j-\frac{1}{2}} := \frac{\mathcal{A}^n_{p,i,j}+\mathcal{A}^n_{p,i,j-1}}{2}, \\ \mathcal{A}^n_{p,i+\frac{1}{2},j} := \frac{\mathcal{A}^n_{p,i,j}+\mathcal{A}^n_{p,i+1,j}}{2}, \\ \mathcal{A}^n_{p,i-\frac{1}{2},j} := \frac{\mathcal{A}^n_{p,i,j}+\mathcal{A}^n_{p,i-1,j}}{2}. \end{cases} \tag{A.1}$$

Further, the central differences are given by

$$\begin{cases} [c_{p,y}]^n_{i,j+\frac{1}{2}} := \frac{c^n_{p,i,j+1}-c^n_{p,i,j}}{\Delta y}, \\ [c_{p,y}]^n_{i,j-\frac{1}{2}} := \frac{c^n_{p,i,j}-c^n_{p,i,j-1}}{\Delta y}, \\ [c_{p,x}]^n_{i+\frac{1}{2},j} := \frac{c^n_{p,i+1,j}-c^n_{p,i,j}}{\Delta x}, \\ [c_{p,x}]^n_{i-\frac{1}{2},j} := \frac{c^n_{p,i,j}-c^n_{p,i-1,j}}{\Delta x}. \end{cases} \tag{A.2}$$

$$\begin{aligned} D_p\Delta c_p - \nabla\cdot[c_p\mathcal{A}_p(x,t,\mathbf{u}(\cdot,t))] = \\ \simeq \frac{D_p[c_{p,x}]^n_{i+\frac{1}{2},j} - D_p[c_{p,x}]^n_{i-\frac{1}{2},j} - c^n_{p,i+\frac{1}{2},j}\cdot\mathcal{A}^n_{p,i+\frac{1}{2},j} + c^n_{p,i-\frac{1}{2},j}\cdot\mathcal{A}^n_{p,i-\frac{1}{2},j}}{\Delta x} \\ + \frac{D_p[c_{p,y}]^n_{i,j+\frac{1}{2}} - D_p[c_{p,y}]^n_{i,j-\frac{1}{2}} - c^n_{p,i,j+\frac{1}{2}}\cdot\mathcal{A}^n_{p,i,j+\frac{1}{2}} + c^n_{p,i,j-\frac{1}{2}}\cdot\mathcal{A}^n_{p,i,j-\frac{1}{2}}}{\Delta y}. \end{aligned} \tag{A.3}$$

Next, we shift our attention to the numerical approach of the adhesive flux $\mathcal{A}_p$ (that explores the effects of cell-cell and cell-matrix adhesion of cancer cells subpopulations). We carry out these computations off-grid by decomposing the sensing region $\mathbf{B}(x,R)$ in $s\in\mathbb{N}^*$ annuli centred at x (with the inner most central circle being of numerically negligible size), and for each annuli $k\in\{1,\dots,s\}$ (counted from the inner most to the outer most), we consider a radial decomposition of this in $2^{h+(k-1)}$ uniformly distributed radial sectors (with $h\in\mathbb{N}$ fixed), as illustrated in the right side of Figure 6 (for $h=2$). This leads to a collection of sectors $\{\mathcal{S}_\nu\}_{\nu=1,N_s}$, where

$$N_s := \sum_{k=1}^{s} 2^{h+(k-1)}.$$

Then for each annulus sector $\mathcal{S}_\nu$, we evaluate the total primary cancer cell population c_1, the total mutated cancer cell population c_2 and the total ECM mass distributed on $\mathcal{S}_\nu$ that are given by

$$\omega_{\mathcal{S}_\nu,c_1} := \frac{1}{\lambda(\mathcal{S}_\nu)}\int_{\mathcal{S}_\nu} c_1(x,t)\,dx, \quad \omega_{\mathcal{S}_\nu,c_2} := \frac{1}{\lambda(\mathcal{S}_\nu)}\int_{\mathcal{S}_\nu} c_2(x,t)\,dx, \quad \text{and} \quad \omega_{\mathcal{S}_\nu,v} := \frac{1}{\lambda(\mathcal{S}_\nu)}\int_{\mathcal{S}_\nu} v(x,t)\,dx. \tag{A.4}$$

Further, on each sector $\mathcal{S}_\nu$, we consider the off-grid barycentres of each annulus sector by $\mathbf{b}_{\mathcal{S}_\nu}$ and the values of each of the three densities at these barycentres, namely $c_1(\mathbf{b}_{\mathcal{S}_\nu},t)$ $c_2(\mathbf{b}_{\mathcal{S}_\nu},t)$ and $v(\mathbf{b}_{\mathcal{S}_\nu},t)$, are obtained via interpolation with bilinear shape functions on the grid rectangles $\{y^i_{\mathbf{b}_{\mathcal{S}_\nu}}\}_{i=1,4}$ that contain $\mathbf{b}_{\mathcal{S}_\nu}$. Therefore, the approximation of the adhesion integral at each instance of time $t>0$ is given by

$$\mathcal{A}_p(x,t,\mathbf{u}(\cdot,t)) = \sum_{\nu=1}^{s2^m} \frac{\lambda(\mathcal{S}_\nu)}{R}\mathbf{n}(\mathbf{b}_{\mathcal{S}_\nu})\cdot\mathcal{K}(\mathbf{b}_{\mathcal{S}_\nu})\,g_p(\tilde{\mathbf{u}}(\mathbf{b}_{\mathcal{S}_\nu},t)), \tag{A.5}$$

where

$$\tilde{\mathbf{u}}(\mathbf{b}_{\mathcal{S}_\nu},t) := [\omega_{\mathcal{S}_\nu,c_1}(t), \omega_{\mathcal{S}_\nu,c_2}(t), \omega_{\mathcal{S}_\nu,v}(t)]^T,$$

and $g_p(\tilde{\mathbf{u}}(\mathbf{b}_{\mathcal{S}_\nu},t))$ is the $p-$th component of

$$g(\tilde{\mathbf{u}}(\mathbf{b}_{\mathcal{S}_\nu},t)) = \left[\mathbf{S}_{cc}[\omega_{\mathcal{S}_\nu,c_1}(t), \omega_{\mathcal{S}_\nu,c_2}(t)]^T + \omega_{\mathcal{S}_\nu,v}(t)\mathbf{S}_{cv}\right]\cdot(K_c - \mathbf{c}(\mathbf{b}_{\mathcal{S}_\nu},t) - v(\mathbf{b}_{\mathcal{S}_\nu},t))^+.$$

B. Parameters used in computations

For all the cancer cells mutation laws reconstructions considered in this work, we use the non-dimensional parameter set specified in Table 1.

Table 1. Summary of the parameter values that have been used for the two local and nonlocal sub-population of cancer cells.

Parameter	Value	Description	Reference
D_1	0.00675	diffusion of primary tumour	[11]
D_2	0.00675	diffusion of secondary tumour	[11]
η_1	2.85×10^{-2}	haptotaxis to ECM from c_1	[37]
η_2	2.85×10^{-2}	haptotaxis to ECM from c_2	[37]
μ_c	0.25	proliferation of tumour cells c	[38]
K_c	1	tissue carrying capacity	[16]
ρ	2	ECM degradation coefficient	[39]
μ_v	0.40	ECM remodelling coefficient	[16]
$\mathbf{S}_{cc}$	$\begin{pmatrix} 0.5 & 0 \\ 0 & 0.3 \end{pmatrix}$	cell-cell adhesion function	[11]
$\mathbf{S}_{cv}$	$\begin{pmatrix} 0.1 & 0 \\ 0 & 0.5 \end{pmatrix}$	cell matrix adhesion function	[11]
$t_{1,2}$	10	time initiation for mutations	[38, 40]
t_s	3	time-steepness coefficient	[38]
δ_0	0.3	mutation from primary tumour	[38, 40]
Δx	0.03125	discretisation step size for $\mathcal{G}_\Omega$	[16]
Δt	10^{-3}	time step size	[16]
$\Delta\eta$	0.0625	mesh size used for $\mathcal{G}_M^1$ and $\mathcal{G}_M^2$	Estimated