1 Definining the Cellular Automaton

We develop a deterministic, spatio-temporal cell-based model to study cancer invasion in a "sandwich" collagen model with a low-density interface. In particular we define a continuous cellular automaton (CA) model. CAs have been used extensively to study biologic processes ranging from the patterns of seashells [2] to the characterization of different modes of metastatic tumor invasion [3]. The CA is defined on a d-dimensional regular grid $\mathcal{G} \subset \mathbb{R}^d$. Each node is connected to its b nearest neighbors with by a set of unit vectors $\mathbf{c_i}$. The distance between nearest neighbors is defined by the grid spacing ϵ. For a two-dimensional grid the unit vectors are defined as:

$$c_i = \begin{pmatrix} \cos\frac{2\pi i}{b} \\ \cos\frac{2\pi i}{b} \end{pmatrix} \quad i = 0, 1, ..., b-1 \tag{1.1}$$

Each node $\mathbf{r}$ is defined by the local cell density $n(\mathbf{r})$ and the local ECM density $f(\mathbf{r})$, both of which take values in the closed interval from 0 to 1 (classification as *continuous* cellular automaton [4]). Biologically, these density values can be interpreted as probability densities providing a measure of relative likelihood of a cell or an ECM molecule to be located at a specific point in space. The state of a node $\mathbf{r}$ is therefore defined by a set

$$\mathcal{S}(\mathbf{r}) = (n(\mathbf{r}), f(\mathbf{r})) \in [0, 1]^2 \tag{1.2}$$

The dynamics of a node $\mathbf{r}$ is determined by the nodes $\mathbf{r'}$ in its interaction neighborhood $\mathcal{N}(\mathbf{r})$:

$$\mathcal{N}(\mathbf{r}) = \{\mathbf{r'} \in \mathcal{G}, d(\mathbf{r}, \mathbf{r'}) \leq R\} \tag{1.3}$$

Where d is the euclidean distance and R is set to the grid spacing ϵ.
The time evolution is defined in discrete time steps $k \in \mathbb{N}$. In each time step a number of deterministic interaction rules are applied updating the states $\mathcal{S}(\mathbf{r})$ and thereby the local cell and ECM densities synchronously.

1.1 Evolution of Local ECM Density

As described the ECM is modeled as a scalar density field were the local ECM density is given by $f(\mathbf{r})$. The temporal evolution can be defined by a discrete ODE updating the ECM density of each node in one discrete time step $k \to k + 1$

$$f(\mathbf{r}, k+1) \coloneqq f(\mathbf{r}, k) - d_1 f(\mathbf{r}, k) n(\mathbf{r}, k) - d_2 \sum_{i=1}^{b} f(\mathbf{r}, k) n(\mathbf{r} + \mathbf{c_i}, k) \tag{1.4}$$

were the assumptions is made that ECM is static meaning there is no random diffusion or other migratory behavior of ECM. Secondly, in the case of uninhibited proteolytic activity ECM gets

removed within the node with a the rate $d_1 \geq 0$ and its neighboring nodes with the rate $d_2 \geq 0$ ($d_1 \gg d_2$). This models the degradation process of ECM by cells, which has been explained to be one of the core principles of ECM remodeling within the tumor growth process [5][6]. Degradation is modeled as a linear increasing function of ECM and cell densities. In the case of non-proteolytic migration were MMPs are inhibited the degradation constant $d_{1,inhib}$ and $d_{2,inhib}$ are set to zero resulting in a time independent local ECM density

$$d_1 = d_2 = 0 \rightarrow \Delta_k f(\mathbf{r}) = f(\mathbf{r}, k+1) - f(\mathbf{r}, k) = 0 \tag{1.5}$$

1.2 Evolution of the Local Cell Density

As the ECM cells are also modeled as a scalar fields were each node is assigned a cell density $n(\mathbf{r})$. However, in comparison to the ECM cells are not static meaning the temporal evolution has to consider migration as well as cell proliferation. The update rule for the cell densities is defined as

$$n(\mathbf{r}, k+1) = 2^{1/\tau} n(\mathbf{r}, k) + M(\mathbf{r}, k) \tag{1.6}$$

with the mitosis rate of $2^{1/\tau}$, where τ is the doubling time in units of one time step, guaranteeing that after τ time steps the cell density has doubled. The migration function $M(\mathbf{r}, k)$ is divided into a inward M_+ and outward migration M_- expression

$$M(\mathbf{r}, k) = M(\mathbf{r}, k)_+ - M(\mathbf{r}, k)_- \tag{1.7}$$

with

$$M(\mathbf{r}, k)_+ = \mu \exp \frac{-f(\mathbf{r}, k)}{f_c} \sum_{i=1}^{k} n(\mathbf{r} + \mathbf{c_i}, k) \sum_{j=1}^{k} f(\mathbf{r} + \mathbf{c_i} + \mathbf{c_j}, k) \tag{1.8}$$

$$M(\mathbf{r}, k)_- = \mu n(\mathbf{r}, k) \sum_{i=1}^{k} \exp \frac{-f(\mathbf{r} + \mathbf{c_i}, k)}{f_c} \sum_{j=1}^{k} f(\mathbf{r} + \mathbf{c_j}, k) \tag{1.9}$$

Assuming homogeneous ECM $f(\mathbf{r}) = f$ and cell density $n(\mathbf{r}) = n$ the migration function M can be written as

$$M \approx nf \exp \frac{-f}{fc} \tag{1.10}$$

As the function M determines the increase (or decrease) of cell density in a time step due to migration, it can be also be understood as a measure for migration speed. A graphic representation of the behavior is shown in Fig.1.1

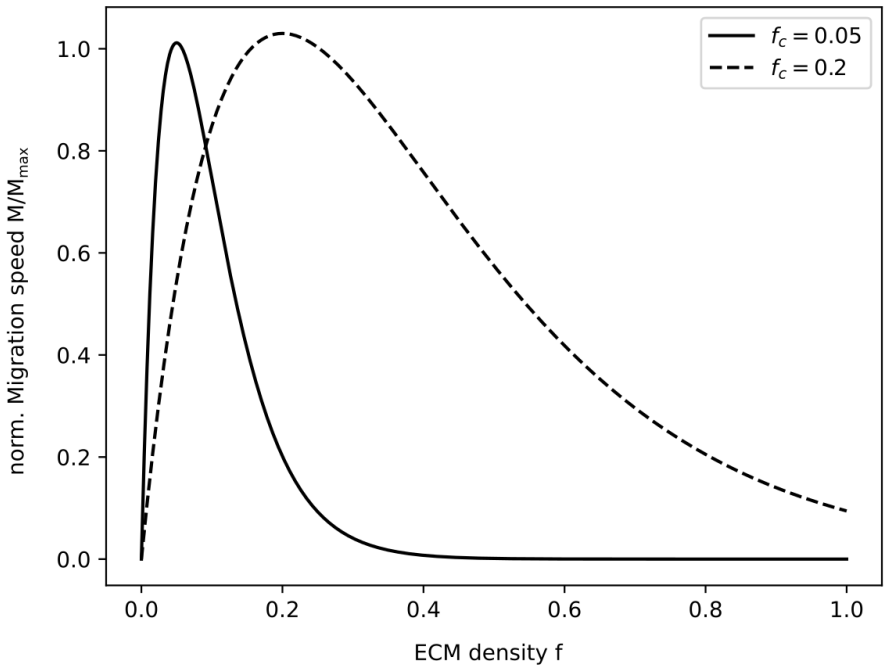

Figure 1.1: **Migration speed of cells in the collagen-collagen culture is strongly dependent on the micro environment.** Depicted is the normalized migration speed of cells M/M_{max} depending on the ECM density in the neighborhood. The migration function M takes the form of $M = n_0 f \exp \frac{-f}{f_c}$. The cell density n_0 is set to 1 and the ECM density is varied between 0 and 1. Shown are two example curves for $f_c = \{0.05, 0.2\}$. The parameter f_c determines when cell migration speed is maximal due to guidance by the critical ECM density f_c.

The underlying biologic principle of this migration function is the combination of two opposing effects which lead to an initial increase of cell speed and a subsequent exponential decrease with increasing ECM density. This is explained by the fact that cell migration is relying on cell-ECM adhesions which are more often achieved in a denser ECM environment, however with increasing ECM density the resisting force due to steric effects between cells and the ECM also increases [7]. This results in the functional behavior depicted in Fig.1.1, which is the simplest description of cell migration in an ECM environment. As for example a model were migration speed just linearly decreases with surrounding ECM density neglects contact guidance and is therefore not biologically relevant for examining tumor invasion processes.

2 Simulating Cell Migration into the Collagen-Collagen interface

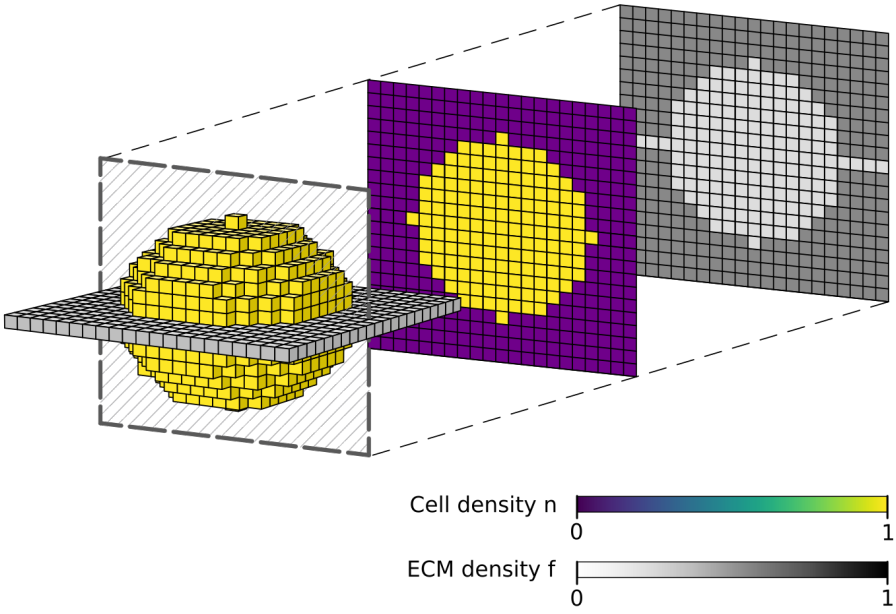

Figure 2.1: **Caricature of the inital ECM and cell densities of the *Sandwich* Model** To simulate the in vitro experiments, the 3 dimensional grid (grid spacing of $10\mu m$) was initiated with a spherical volume of high cell density n_0 (yellow) surrounded by with high ECM density $f_{3D,0}$. Cell and ECM densities are indicated by their respective color codes. The radius of the spheroid was set to $200\mu m$ matching the experimental condition. Additionally, a plane of low density ECM $f_{int,0}$ was implemented splitting the ECM into 3 regions creating the characteristic "sandwich-like" formation. Furthermore the ECM density at nodes $\mathbf{r}$ corresponding to the initial spheroid of cells where set to 0, as it was assumed that the initial cell bulk quickly degraded the local ECM.

In the experiment the total numbers of cells which migrated into the *random* region and the total number of cell migrating into the interface were measured. The exact number of cells can not be obtained from the model, as only the cell densities are tracked for each node and time step (Fig.2). However, as the local cell density can be understood as relative likelihood a cell to be found at this point in space. The sum over all densities of a given grid gives an estimate how many cells are expected to be found within the space. Applying this notion and setting the origin to the midpoint of the initial sphere the total quantity of migrating cells is given by:

$$N_{3d} \coloneqq \sum_{\mathbf{r} \in G_{3d}} n(\mathbf{r}) \tag{2.1}$$

$$G_{3d} \coloneqq \{\mathbf{r} \in \mathcal{G},\, d(\mathbf{r}) \geq R_0\} \tag{2.2}$$

where d is the euclidean distance and R_0 the initial radius of the spheroid. To distinguish between migration into the random region and the interface it is useful to define regions in space which are correlated with random migration and interface migration. As the interface is defined as the xy-plane the volume which is unaffected by interface migration is given by the pyramid comprised of the nodes:

$$G_{random} \coloneqq \{(x, y, z)^T \in G,\, z \geq 0 \wedge d(x, y) \in [-z, z]\} \tag{2.3}$$

It is further useful to only take into account nodes outside the initial sphere. The constrained $\tilde{G}_{random}$ can therefore be defined as the intersection:

$$\tilde{G}_{random} \coloneqq G_{random} \cap G_{3d} \tag{2.4}$$

Now the total number of cells which migrated into the *random* regions is defined as

$$N_{random} \coloneqq \sum_{\mathbf{r} \in \tilde{G}_{random}} n(\mathbf{r}) \tag{2.5}$$

The number of cell migrating into the interface can therefore be calculated by the difference between N_{tot} and N_{random}:

$$N_{int} = N_{tot} - 6N_{random} \tag{2.6}$$

with a factor 6 (in 3D), which scales the volume of G_{random} to the whole grid, as can be seen by

$$|G_{3d}| = 6|G_{random}| \tag{2.7}$$

where $|\cdot|$ denote the respective cardinalities. From this absolute migration numbers the interface migration ratio is given by:

$$N_{int,rat} \coloneqq \frac{N_{int}}{N_{tot}} \tag{2.8}$$

The introduced parameters for total $N_{int,tot}$ and relative $N_{rat,tot}$ migration into the interface can now be used compare and fit the computational model to the *in vitro* cell data.

3 Fitting the Model to the Experimental Data

The experiments showed that for both the "high" (6mg/ml) as well as the "low" (2mg/ml) density collagen cell migration was observed predominantly into the interface in both the presence presence and absence of MMP inhibitors. In the non-proteolytic experiments the ratio of interface migration was highest for both condition especially in the case of "high" density collagen where migration into the *random* 3D ECM lattice was almost completely suppressed. Secondly, the effect of proteolysis decreased the ratio of interface to total migration but increased the absolute cell migration into the interface. Thereby defining both guidance by the ECM geometry as well as matrix degradation as determinant of cell patterning. The introduced model output values of $N_{int,tot}$ and $N_{int,rat}$ were used and compared to the experimental values of total number of cells in the interface after 48 hours as well as the ratio of number of cells in the interface to total number of emigrated cells (data is shown in Fig.4 of [1]). First simulation outputs without proteolysis after 100 timesteps (d=0) of all combinations of initial ECM densities ranging between 0 and 1 in both the interface $f_{int,0}$ and the *random* 3D region $f_{3D,0}$ were tested and compared to the non-proteolytic experimental migration data. This led to a range of possible values for both $\tilde{f}_{int,0} \in F_{int,d=0}$ and $\tilde{f}_{3D,0} \in F_{3D,d=0}$ with

$$F_{d=0} \subset \{x \,|\, 0 \leq x \leq 1\} \tag{3.1}$$

which were than used as new ranges for the parameter scan for the proteolytic case where again all combinations of $\tilde{f}_{int,0}$, $\tilde{f}_{3D,0}$ and additionally a non-zero degradation constant $d \in (0, 1]$ were tested. Subsequently, the simulation outputs of relative and absolute migration into the interface (Fig.3.2) were again compared to the experimental data now without inhibition of proteolysis. The resulting parameter ranges - fitting both the proteolytic and the non-proteolytic condition - for $f_{int,0}$, $f_{3D,0}$ and d as well as other simulation parameters are listed in Table 1.1.
In Fig.3.1 a specific model realization of the temporal spatial evolution of cell density is shown for both the 6mg/ml and 2mg/ml conditions.

Table 3.1: *Sandwich* model parameters

Parameter	Description	2mg/ml	6mg/ml
t	number of time steps	100	100
δt	length of one time step	0.5h	0.5h
τ	doubling time of MV3 cells	24h [8]	24h [8]
μ	migration const.	0.2	0.2
n_0	initial cell density	1	1
f_c	critical ECM density	0.1	0.1
$f_{int,0}$	initial ECM density in the interface	[0.2, 0.32]	[0.08, 0.12]
$f_{3D,0}$	initial ECM density in the random 3D region	[0.46, 0.52]	[0.74, 1]
d_1	degradation const. of ECM in the respective node	[0.04, 0.08]	[0.12, 0.18]
d_2	degradation const. of ECM in neigbouring nodes	$0.1d_1$	$0.1d_1$

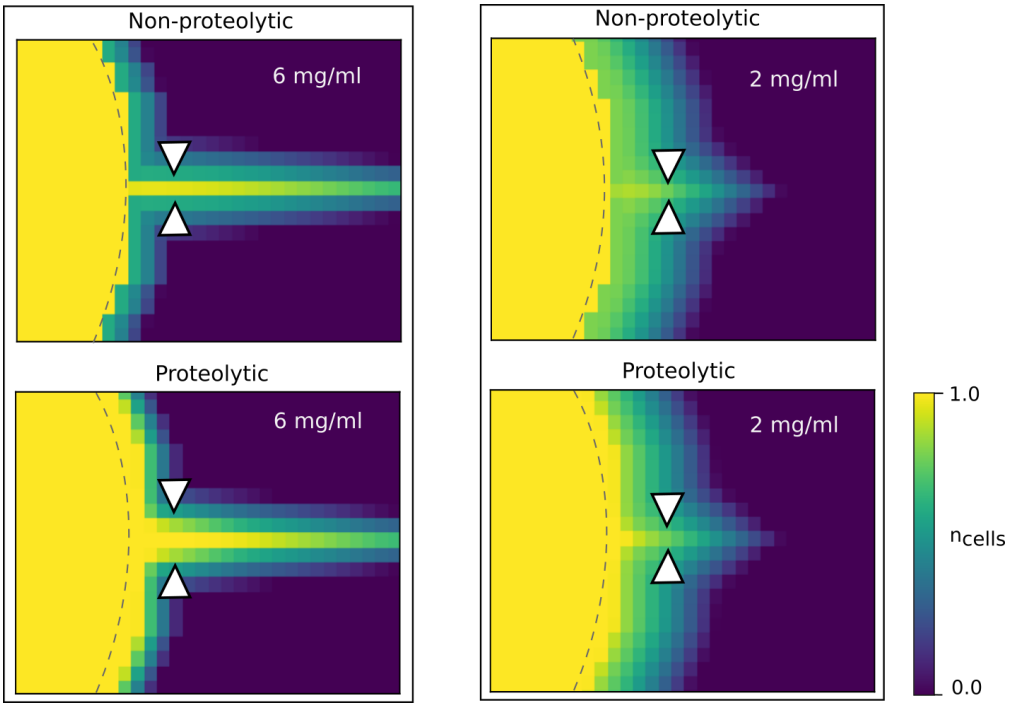

Figure 3.1: **Example *in silico* realization of tumor growth into a low density interface** Shown are zoomed in snapshots (xz-plane) on the interface close to the initial spheroid (grey dotted line) for the cell densities (as indicated by the color code) after 100 discrete simulation steps, were the spheroid already migrated into the interface (marked by the arrow heads) and the *random* region. The two different collagens (6mg/ml and 2mg/ml) are shown for both the non-proteolytic ($d = 0$) and the proteolytic condition ($d > 0$). Specifically, in the case of the 6mg/ml collagen (left) cells migrated strongly into the interface while migration was minimal. If proteolysis was not inhibited a widening of the interface close to the initial spheroid was observed in the high density collagen (also visible to a much lesser extent in the low density collagen). Simulation parameters for the 6mg/ml were $f_{3D,0} = 0.9$, $f_{int,0} = 0.1$ and $d = 0.15$. For the 2mg/ml collagen $f_{3D,0} = 0.46$, $f_{int,0} = 0.32$ and $d = 0.08$ were chosen. All other simulation parameters were taken from Table 1.1.

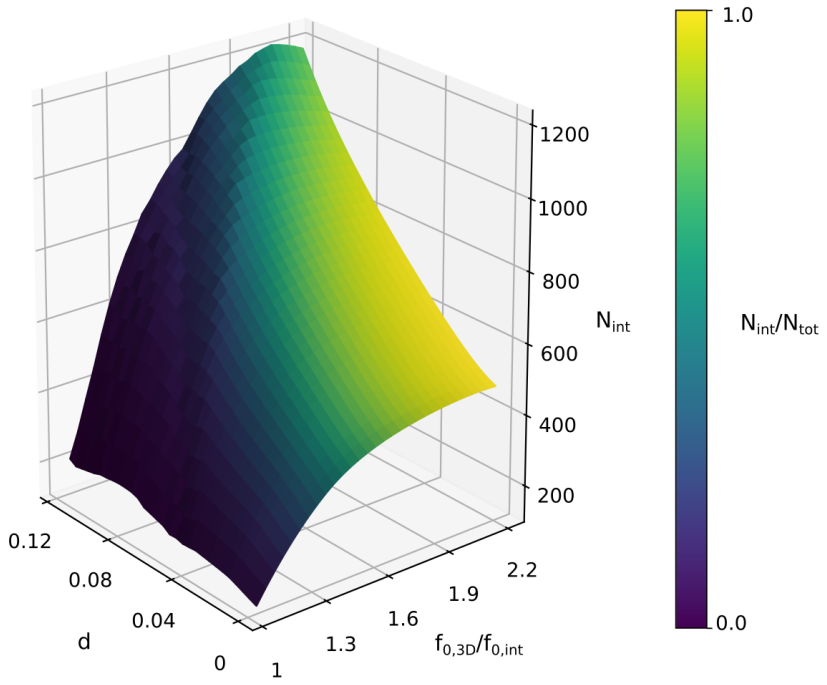

Figure 3.2: **Interface migration increases with higher 3D ECM density and degradation rate** Shown is the number of cells in the interface after 100 time steps against the ratio of initial ECM densities $\frac{f_{0,3D}}{f_{0,int}}$ and the degradation constant d. The relative interface migration compared to the total migration N_{int}/N_{tot} is indicated by the color code. Interface migration increases both absolutely as well as relatively with increasing ECM density in the 3D region. Furthermore for high $f_{0,3D}$ proteolysis increases the number of cells in the interface while the percentage of interface migration remains high. The simulation parameters were chosen in a way that the experimental values for N_{int} and N_{int}/N_{tot} for the 2mg/ml are depicted. The initial ECM density in the interface was set to 0.26. Both $f_{0,3D}$ as well as d were varied in 30 discrete steps with constant step size $f_{0,3D} \in [0.26, 0.6]$ and $d \in [0, 0.12]$. The simulation time was 100 time steps. All other simulation parameters were taken from Table 1.1. The noise for high degradation rate and low $f_{0,3D}$ is due to numerical rounding errors where ECM densities are close to 0.

4 Results and Discussion

The presented CA model of interacting cell and ECM densities is able to reproduce relative and absolute cell numbers of emigrating cells into a collagen-collagen interface in the case of proteolysis and with inhibition of proteolysis. Importantly, the model was only able to reproduce the preferential migration mode into the interface, because cell migration was modeled in a way that high neighboring ECM density increases cell speed (contact guidance). A model were migration speed just linearly decreased with surrounding ECM density (no contact guidance) $M \propto (1-f)$ could not reproduce the high cell numbers in the interface for both conditions. Further indicating that ECM geometry gives critical guidance cues for cell migration. In addition the model was able to predict that cells degraded the high density collagen (6 mg/ml) collagen gel with a degradation rate (mean $d_{6mg/ml} = 0.15$) around 2.5 times faster than the density low density (2 mg/ml) collagen (mean $d_{2mg/ml} = 0.06$), which is in good agreement with previous experimental findings [9].
However, the model failed in reproducing different invasion phenotypes of single-cell and strand-like invasion into the randomly organized 3D matrix. This is due to the fact that the described model uses a density approach to model cell invasion. In comparison to "normal" cellular automaton models, were cells are modeled as discrete agents and can be tracked, the states in the *sandwich* model are continuous in the range of 0 to 1. Therefore it can be classified as a *continuous* automaton [4], with a discrete lattice. The definition of continuous state variables as probability densities has the main benefit that the dynamic of the system can by described by a set of deterministic update rules which drastically reduces the amount of computation needed. This is especially true for an increasing amount of model parameters which have to be fit to the experimental data (here initial ECM densities $f_{3D,0}$, $f_{int,0}$ and degradation rate d as computational time needed increases exponentially with the number parameters which have to be varied and evaluated.
To reproduce the observed plasticity of invasion phenotypes an agent-based model is needed incorporating single cell interaction with their micro environment.

Bibliography

[1] L. Beunk, S. van Helvert, B. Bekker, L. Ran, R. Kang, T. Paulat, S. Syga, A. Deutsch, P. Friedl, and K. Wolf, "Extracellular matrix guidance determines proteolytic and non-proteolytic cancer cell patterning," *bioRxiv*, 2022.

[2] S. Coombes, "The geometry and pigmentation of seashells," *Dep. of Math. Sciences, University of Nottingham, Nottingham*, 2009.

[3] O. Ilina, P. G. Gritsenko, S. Syga, J. Lippoldt, C. A. La Porta, O. Chepizhko, S. Grosser, M. Vullings, G.-J. Bakker, J. Starruß, *et al.*, "Cell–cell adhesion and 3d matrix confinement determine jamming transitions in breast cancer invasion," *Nat. Cell Biology*, vol. 22, no. 9, pp. 1103–1115, 2020.

[4] B. MacLennan, *Continuous spatial automata*. University of Tennessee, Computer Science Department, 1990.

[5] C. Bonnans, J. Chou, and Z. Werb, "Remodelling the extracellular matrix in development and disease," *Nat. Rev. Mol. Cell Biology*, vol. 15, no. 12, pp. 786–801, 2014.

[6] J. Winkler, A. Abisoye-Ogunniyan, K. J. Metcalf, and Z. Werb, "Concepts of extracellular matrix remodelling in tumour progression and metastasis," *Nat. Commun.*, vol. 11, no. 1, pp. 1–19, 2020.

[7] J. Zhu and A. Mogilner, "Comparison of cell migration mechanical strategies in three-dimensional matrices: a computational study," *Int. Focus*, vol. 6, no. 5, p. 20160040, 2016.

[8] G. N. van Muijen, K. F. Jansen, I. M. Cornelissen, D. F. Smeets, J. L. Beck, and D. J. Ruiter, "Establishment and characterization of a human melanoma cell line (mv3) which is highly metastatic in nude mice," *Int. J. of Cancer*, vol. 48, no. 1, pp. 85–91, 1991.

[9] E. Infante, A. Castagnino, R. Ferrari, P. Monteiro, S. Agüera-Gonzalez, P. Paul-Gilloteaux, M. J. Domingues, P. Maiuri, M. Raab, C. M. Shanahan, *et al.*, "Linc complex-lis1 interplay controls mt1-mmp matrix digest-on-demand response for confined tumor cell migration," *Nat. Commun.*, vol. 9, no. 1, pp. 1–13, 2018.