MASTER DISSERTATION

CompuCell3D Model of Cell Migration Reproduces Chemotaxis

Pedro C. Dal-Castel[1], Gilberto L. Thomas[1],

Gabriel C. Perrone[1], and Rita M.C. de Almeida[1,2,3*]

[1]*Instituto de Física, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil*

[2]*Instituto Nacional de Ciência e Tecnologia: Sistemas Complexos,*

Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil and

[3]*Programa de Pós-Graduação em Bioinformática, Instituto Metrópole Digital,*

Universidade Federal do Rio Grande do Norte, Natal, RN, Brazil

(Dated: June 2023)

Abstract

The introduction and conclusion of the following dissertation serve to support and summarize our research entitled "CompuCell3D Model of Cell Migration Reproduces Chemotaxis." In this study, we created a CompuCell3D simulation of single cell chemotaxis, a biological phenomena in which cells move in response to environmental chemical cues. We also developed an analysis scheme to analyze recordings of center of mass and polarization over time to characterize cell dynamics and kinetics. Aiming at individuals with intermediate modeling experience who lack specific understanding in the field, we offer the relevant biology, mathematics, and computational foundation in order to adequately prepare the reader. In the first topic, we discuss the biological cell and its capacity to migrate, discussing both the significance of this capacity for survival and the underlying biochemical mechanism. Second, we explore a few computational and mathematical models of cell migration, focusing on a brand-new analytical model called the Anisotropic Ornstein-Uhlenbeck Process, which treats polarization in its stochastic differential equations. Finally, we go over CompuCell3D's functionality in detail and provide a real-world example for readers to try out (needs access to a computer with Windows installed). Our research on single cell movement aims to completely characterize chemotaxis and offer tools that may be used to analyze experimental data, provided that cell polarization is measured. We discuss the significance of cell polarization measurements and the proper way to handle the issue of cell velocity when short time scales exhibit cell diffusive behavior. We suggest a procedure for measuring chemotactic efficiency as well as a way to discriminate between cell reorientation and cell drift speed modulation as chemotactic response modalities. Our simulation serves as the basis for upcoming collective migration models and may be utilized to investigate the role of particular types of white blood cells during innate immune response.

───────
* pdalcastel@gmail.com; glt@if.ufrgs.br; gabriel.perrone@if.ufrgs.br; rita@if.ufrgs.br

THEORETICAL INTRODUCTION

Here we aim at a basic and elaborate introduction to biological and mathematical concepts used in the study of single cell migration and chemotaxis. We will also explain in detail how Cellular Potts Model (CPM) [1–3] works, and review mathematical models of cell migration such as Langevin equation and the Anisotropic Ornstein-Uhlenbeck Process. By the end of this theoretical introduction, the reader will be ready to understand the proposed model and results presented in our paper. This introduction is divided as follows: **A.** overview of a biological cell, how it moves, and how it interacts with the environment; **B.** computational models in cellular biology, and mathematical models of single cell migration using stochastic differential equations; **C.** the Cellular Potts Model (CPM) and its implementation in CompuCell3D (CC3D) with a practical example of a real cellular biology problem that the reader can follow through as a first CompuCell3D experience.

A.   From Proteins to Motility: How Do Cells Orient and Migrate?

The basic structure of a cell consists in **1)** a membrane separating interior (cytoplasm) from exterior (environment), **2)** specific molecules carrying information about cell behaviors and instructions on self replication (RNA/DNA), **3)** other molecules in the cytoplasm that allow cell **metabolism** to run, such as nutrients, proteins, ions etc. (carbohydrates, ATP, enzymes, Ca^{2+} ...).

Despite the fact that the word "metabolism" literally means "change", in cellular biology, it refers to all chemical processes that the cell sustains in order to survive, proliferate, or die. Metabolic processes can be divided into anabolism and catabolism. Catabolism is the breakdown of molecules. The most notable catabolism mechanism is probably the breakdown of absorbed molecules to release energy. The additional energy can then be used for anabolic tasks like maintenance and building.Although processes like diffusion, heat transfer, and mechanical deformations are excluded from this definition, it is important to note that cells can alter their metabolism to interfere with each of these processes. For instance, the cytoskeleton can control mechanical deformations, and ion pumps and channels can control ion diffusion. Although the transition from alive to dead (such as apoptosis) is considered metabolism by the preceding definition and may be a kind of catabolism, we define cell

death as the irreversible condition of zero metabolism. Death is not the same as having no metabolism. The expression "irreversible" is crucial. As an example, extremophilic bacteria can cease their metabolism for years in face of extreme desiccation and reversibly recover after hydration [4].

Every cell engages in a number of metabolic activities, including the conversion of nutrients into energy, movement of molecules throughout the cytoplasm, and removal of metabolic wastes. However, sometimes these abilities are insufficient. A bacterium, for example, may need to migrate if nutrients deplete, if temperature or light intensity are not favorable for survival, or if toxic substances are accumulating. To orient their migration direction according to external chemical gradients, some cells sense spacial concentration variations, leading to chemotaxis; other cells sense temporal concentration variations, leading to chemokinesis. Directed migration then take place until cells find better environment conditions. In multicellular organisms, eukaryotic cell migration plays important roles in regulation of tissue development, immunologic response, wound healing and cancer metastasis [5].

Nature has found many solutions for cell locomotion: propulsive engines like flagella, swimming engines like cilia, and stick and pull structures like lamellipodium. Some of these structures can also probe the environment and acquire cues to direct cell movement. Flagellum can behave as a wetness sensor, and Lamellipodium can sense the stiffness of the ECM (extra cellular medium).

Shifting to physiological processes in humans, we now will focus on eukaryotic cells with lamellipodium. They crawl over 2D substrates or through complex 3D ECM by adhering and pulling. Lamellipodium is a flat net of actin filaments inside the cell as you can see in Fig. 1C and Fig. 2. Each actin filament is a chiral, long and stiff structure that polymerize preferentially in one end (barbed end), depolymerizing in the other (pointed end). Actin filaments in the lamellipodium can be limited by capping proteins and can branch via Arp2/3 proteins. A net of these filaments grows everywhere in the cell, creating a global competition for free actin molecules. As filaments branch, more actin is locally used, promoting localized growth, and rising cell membrane tension globally. High demand for actin and high membrane tension disfavor growth in regions of the cell where lamellipodium less is developed. This effect corresponds to a Local Excitation Global Inhibition - LEGI [6–10] dynamics. It explains why lamellipodium can spontaneously polarize and establish a preferential axis, breaking cell's circular symmetry into a (in average) bilateral symmetry.
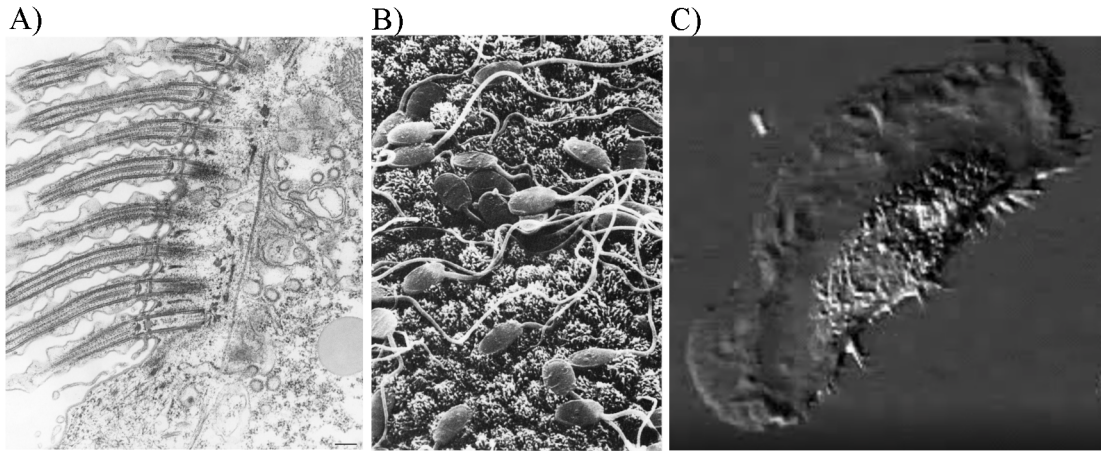
Figure 1. **A**) Transmission electron microscopy image of cilia in a *Paramecium caudatum*. By Richard Allen. Image from CIL – Cell Image Library (cellimagelibrary.org/images/36768 accessed in 02/15/2023). **B**) Scanning electron microscopy image of flagellum in sperm cells. By Don W. Fawcett and David Phillips. Image from CIL – Cell Image Library (cellimagelibrary.org/images/35957, accessed in 02/15/2023). Link to license: creativecommons.org/licenses/by-nc-nd/3.0/. **C**) differential interference contrast microscopy recorded image of lamellipodium in an epidermal cell. By Mark Cooper. Image from CIL – Cell Image Library (cellimagelibrary.org/images/37332, accessed in 02/15/2023). Link to license: creativecommons.org/licenses/by-nc-sa/3.0/.
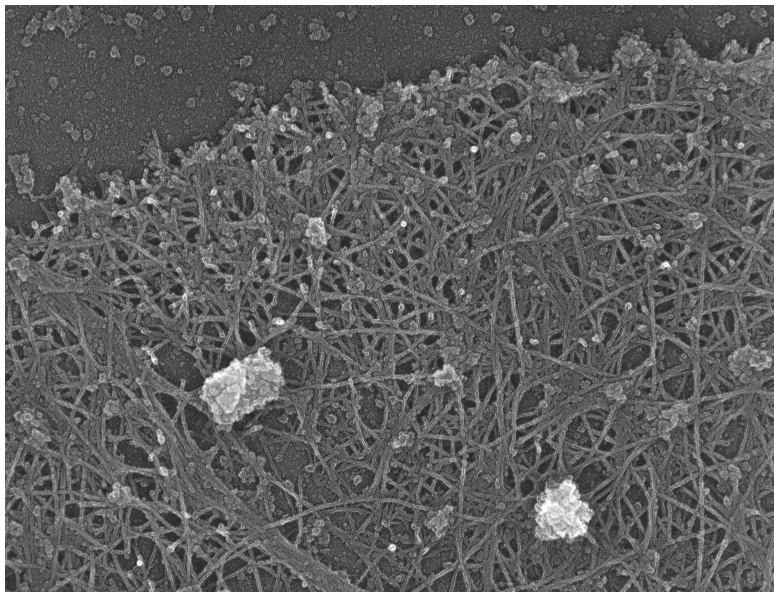

Figure 2. Transmission electron microscopy image of lamellipodium in a vertebrate fibroblast. By Tatyana M. Svitkina and Gary G. Borisy. Image from CIL – Cell Image Library (cellimagelibrary.org/images/24788 accessed in 02/15/2023). Link to license: creativecommons.org/licenses/by-nc-sa/3.0/. You may notice the high density of actin fibers and how frequently they branch. Full article in reference [11].

Lamellipodium adheres to substrate or ECM via integrin transmembrane proteins attached both to actin filaments and to substrate fibers. These adhesion points hold lamellipodium fixed relative to the substrate. As filaments grow and branch in lamellipodium's front and retract at the rear, cell's membrane is pushed and flows forward. As filaments advance into new territory, new adhesions are made between lamellipodium and the substrate, while older adhesions are destroyed in the back. Free actin molecules flow from the rear to the front, allowing the cell to sustain this process and migrate. For reviews on cell migration, check refs. [12, 13]

Extra mechano-chemical mechanisms increase efficiency of migration via lamellipodium, and are sometimes necessary. We list three important processes addressed in ref. [14]: **1)** myosing molecules concentrated at the rear attach to filaments, contract (ATP mediated) to break filaments and adhesions, retracting cell rear and promoting actin flow to the front, **2)** establishment of microtubules along polarization axis create positive feedback with actin polymerization and allow active transport of vesicles carrying regulatory proteins from the rear to the front, and **3)** Rac and Rho GTPases in active/inactive forms distributed inside the cell promote or inhibit polymerization of actin filaments. In Fig. 3, we schematically show how these processes are organized inside the cell.

In addition to fundamental movement, eukaryotic cells contain a membrane full of receptors that can bind to or let in external chemical agents to internalize information from the environment. The contact between the substance and the receptor sets off internal chemical reaction cascades. When perturbation of internal chemical equilibrium is spatially uneven throughout the cell, it can influence cell polarization orientation, which then directs migration, i.e., chemotaxis. The details on how information transduction and cell reorientation take place are specific to cell species, phenotype and external conditions. [15–17]

Many forms of directed migration, other than chemotaxis, have also been observed: electrotaxis (orientation by electrical fields), durotaxis (orientation by substrate stiffness gradients), phototaxis (orientation by light intensity gradient), and others. Understanding directed migration is of great importance for studying wound healing, cancer metastasis, immunologic response and embryo development, as cells rely on environmental cues to achieve large scale organization. [5]

Unfortunately, these processes are complex enough to preclude the use of physics' conservation laws and equilibrium thermodynamics. Biologists often resort to study chemical
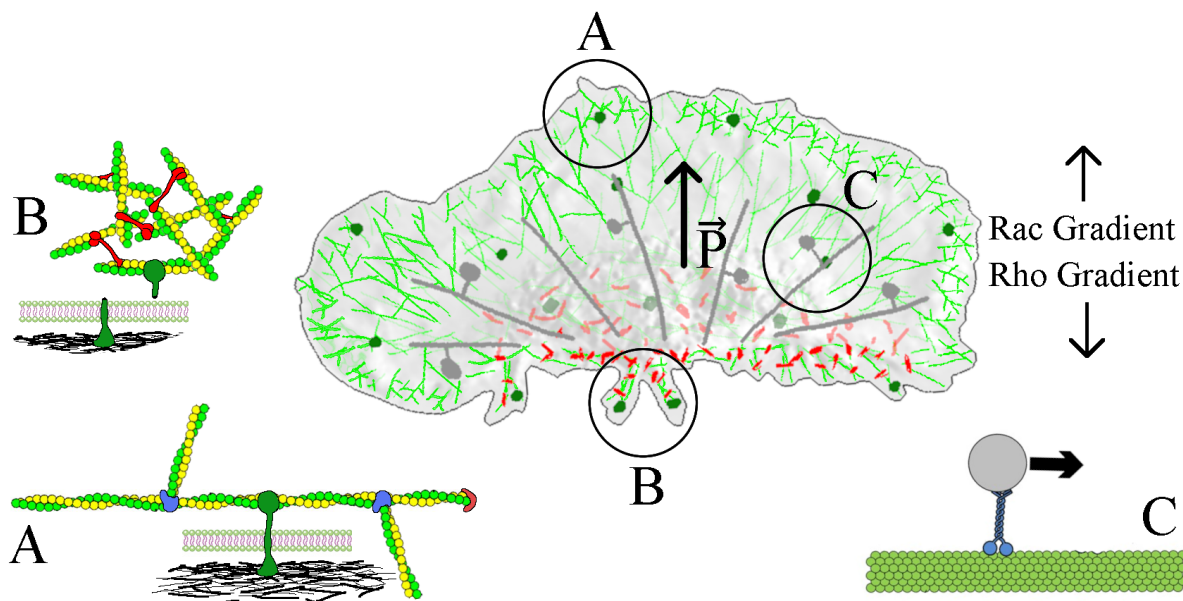
Figure 3. Representation of a cell crawling with lamellipodium and stabilised polarization direction $\vec{P}$. **A**) Polimerization of actin filaments at the front, branching via Arp2/3 and filament's capping. Integrins attach to fibers and to substrate through cell membrane. **B**) Actin fibers detachment from substrate and breaking into smaller pieces via myosin II. **C**) Active transport of vesicles containing actin molecules and other regulatory proteins. Rear contraction helps actin flow to the front, and gradients of active/inactive forms of Rac and Rho GTPases create feedback loops with actin fibers' polymerization dynamics.

pathways, to isolate specific cellular molecules, or merely classify and document observations. Lately, however, modern computation allows us to operate increasingly complex multivariable systems. The use of such computational models enables the proposition of new hypothesis and new experimental setups [18]. We recognize the contribution of computational models as a driving force in the advancement of biology.

B. Models of Cellular Biology: Computational and Mathematical Approaches

Computational models of cellular biology can hardly rely on conservation laws of Physics or on equilibrium thermodynamics. Can you think a way of modelling human behavior (a crowd in a stadium for example) using conservation of energy and momentum? It is almost non sense. Only very small systems and specific cases will allow this type of modelling. When considering large metabolic organisms (cells, tissues and whole animals), complexity preclude the use of classic theories from physics. One possible approach is to consider cells as discrete agents, setting rules of interactions to determine their behavior. This approach

is often referred to as "Game Theory" [19]. Another approach is to consider cells as a continuous field (commonly a density field), employing partial differential equations and boundary conditions to describe the evolution of the system.

Models of continuous cell density fields can use fluid mechanics, partial differential equations (PDE), logistic maps, and Lattice-Boltzmann methods. The great advantage of such models is scalability and the small number of simulation parameters. Moreover, most solving methods are well documented and ready to use in C, C++, Python, Fortran, Julia and other languages.

On the other hand, agent based models provide a bottom-up approach allowing the emergence of collective behavior from individual characteristics. Some examples are: vertex models, where dynamics occur at vertices and edges, with each polygon defining a cell; center models, where cells are self propelled point particles with interacting potentials and equations of motion; phase field models, where each cell is a connected region of high intensity field, that determines system's evolution; and lattice models, like Cellular Automata and Cellular Potts Model, where each cell is a set of pixels that obey a set of rules (Automata) or minimize global energy (Potts). The weaknesses of agent based models are the strengths of continuous models. They often have many parameters and are computationally expensive for large agent numbers. See Fig. 4 for visual representation of some agent based computational models.

Although the center-based approach is the most computationally efficient, it lacks realistic cell properties associated with cell size and shape. Recently, however, Emanuel F. Teixeira and collaborators built a deformable extensive propelled cell (with volume, surface and clear borders) using the center-based approach. See reference [20] for detailed explanation.

Both computational and experimental models require metrics that come from mathematical theories. For example, when studying bacteria colonies, Malthusian growth rate can be used to compare the effect of different parameters over population growth. Malthus model is a simplistic description of population growth, yet it gives clear measurement protocols to characterize computational and experimental systems. What happens if a system deviates from the Malthus model? Well, you will still use Malthus model to demonstrate the discrepancy between prediction and observation, which promotes the sofistication of current theories. Sometimes it is impossible to create a model that can quantitatively predict a system's evolution. In this case, qualitative models can be employed. Although they are
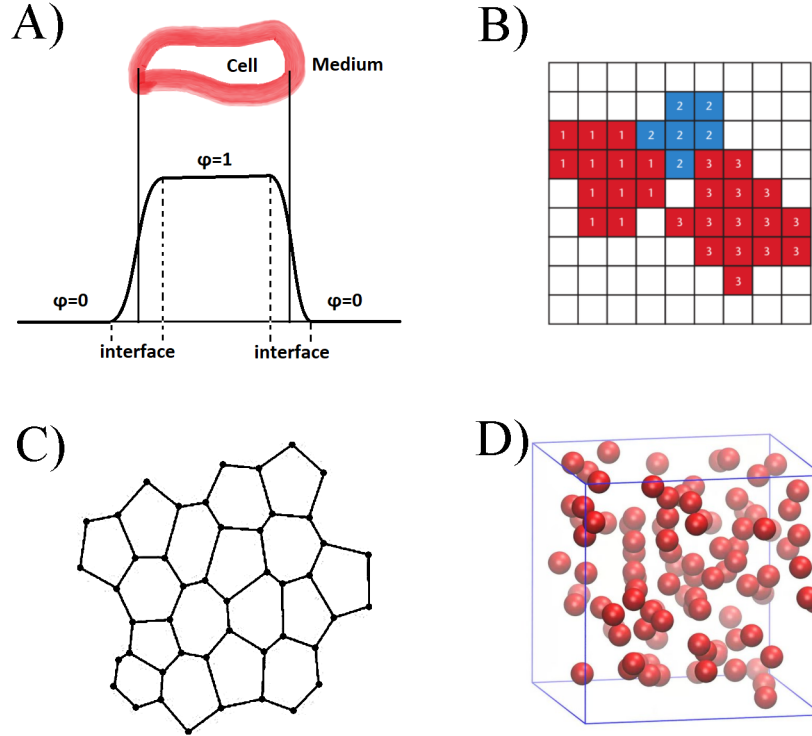
Figure 4. **A**) Representation of a cell in a phase field model. The cell exists where $\phi = 1$. There is a smooth transition between cell and medium ($\phi = 0$), defining an interface. **B**) 2 types of cells (blue and red) and three cells (indexes 1, 2 and 3) represented in the lattice of a CPM simulation. **C**) Representation of a tissue in a vertex model. Each polygon represents a cell. **D**) Representation of cells in a center model simulation. Each cell is a point particle. Often there is short range repulsion between cells, representing cell volume.

not exact, they can give insights, helping people to make decisions, and they can even be used to propose classification criteria. A good example is the evolution of species: even in the absence of a complete quantitative mathematical model, we can organize the species (taxonomy) from our qualitative understanding of natural selection. Theoretical models, whether quantitative or not, are essential for comprehending how a phenomenon operates, for making judgments, and for exploring it sensibly.

Going back to the main topic of this dissertation, i.e. cell migration, the most used mathematical model is Langevin equation for velocity

$$m\frac{dv}{dt} = -\gamma v + \xi \quad . \tag{1}$$

It considers a point particle receiving random collisions from surrounding particles represented by a random force term ξ, and suffering drag from the medium represented by the

viscosity term $-\gamma v$. This equation was originally formulated to solve the Brownian particle problem. In cell migration, ξ is interpreted as a self propelling mechanism that uses energy from cell's metabolism, and $-\gamma v$ is a memory loss due to constant repolymerization of F-actin and interactions with the substrate. In this model, the cell is never in thermodynamic equilibrium, but the system reaches a steady state when the energy input from ξ equals the energy loss to $-\gamma v$.

The analytical solution of Langevin equation for the Mean Square Displacement - MSD

$$\langle |\Delta \vec{r}|^2 \rangle = 2D(\Delta t - P(1 - e^{-\Delta t / P}))$$ (2)

provides important parameters to quantify cell movement: diffusivity and persistence time. Diffusivity distinguishes migration capacity of cells over long periods of time, while persistence time determines cell's ability to sustain a migration speed and direction in the short term. However, this model fails to describe the short term diffusive behavior observed in some experiments, as shown by Thomas and collaborators [21].

To solve this problem, a recent mathematical model of cell migration considers a polarized cell with different dynamics in the parallel and perpendicular directions to polarization vector. This model also considers the effect of membrane and cytoskeleton fluctuations as a noise in cell displacement. The equations of motion

$$\Delta \theta = \int_t^{t+\Delta t} \beta_\perp(t) dt$$ (3)

$$v_\parallel(t + \Delta t) = \left[(1 - \gamma \Delta t) v_\parallel(t) + \int_t^{t+\Delta t} \xi_\parallel(t) dt \right] (\hat{p}(t) \cdot \hat{p}(t + \Delta t))$$ (4)

$$\Delta r_\perp = \int_t^{t+\Delta t} \xi_\perp(t) dt$$ (5)

were proposed by de Almeida and collaborators in 2020 [22]. This model is necessarily 2D since it defines parallel and perpendicular directions to the polarization. The important remarks of this model are: **1)** polarization is a central concept, it is a unit vector whose direction obeys a Wiener process with Gaussian white noise $\beta_\perp$ (Eq. 3), **2)** velocity dynamics takes place in the parallel direction to the polarization, where it obeys a Langevin-like process with an additional memory loss from reorientation, expressed by the scalar product between

10

two successive polarization orientations $\hat{p}(t)$ and $\hat{p}(t + \Delta t)$ (Eq. 4), **3)** in the perpendicular direction, the cell is only affected by fluctuations in membrane and cytoskeleton, which behave as a gaussian white noise displacement $\xi_\perp$ (Eq. 5).

de Almeida and collaborators solved equations 3-5 analytically. They obtained the solution for MSD

$$\langle |\Delta \vec{r}|^2 \rangle = 2D(\Delta t - P(1 - e^{-\Delta t/P})) + \frac{2DS}{1 - S}\Delta t \quad , \tag{6}$$

where the extra term linear in Δt accounts for the diffusive behavior in short time scales. The new parameter S provides information about the time scale below which cell movement is diffusive.

Both models lack a description of chemotactic response, which we need to analyze and interpret data from simulations. In the case of the Langevin equation, simply adding an external force term as

$$m\frac{dv}{dt} = -\gamma v + \xi + F \tag{7}$$

does not reflect the mechanism observed in the experiments. In chemotaxis, the cell is not pulled by the chemical solution, as Eq. 7 implies. In reality, cell's polarization is reoriented, as shown by experiments [15–17, 23].

A more realistic approach would be to consider a restoring torque in polarization direction

$$\Delta\theta = \int_t^{t+\Delta t} \beta_\perp(t)dt - l(\theta(t) - \theta_{eq}) \quad , \tag{8}$$

and add a net positive velocity in the parallel direction to polarization. This new stochastic model has not yet been solved analytically. It is an ongoing project of Guilherme Shoiti and our research group.

We are left without an appropriate theoretical solution. But we can propose a phenomenological MSD curve for chemotaxis. Consider a cell with polarization orientation bias in the x-direction and a net velocity in the direction of polarization. In the steady state, we expect the average displacement in the x direction to increase linearly with time

$$\langle \Delta x'(\Delta t) \rangle = V_T \Delta t \quad , \tag{9}$$

whereas the average displacement of a cell without bias would be $\langle \Delta x(\Delta t) \rangle = 0$. V_T refers to

terminal speed. This concept of terminal speed is used for particles under constant external force plus a viscous force. Despite this not being the case for cells during chemotaxis, we still employ the term "terminal speed", since we also do not expect cells to accelerate indefinitely in such conditions.

Assuming that $x(t)$ and $x'(t)$ obey the same stochastic realization, we can express their difference in displacement as

$$\Delta x'(\Delta t) - \Delta x(\Delta t) = V_T \Delta t \quad . \tag{10}$$

From this equation, we can calculate the mean square displacement in the x-direction, obtaining

$$\langle \Delta x'^2(\Delta t) \rangle - \langle \Delta x'(\Delta t) \rangle^2 = \langle \Delta x^2(\Delta t) \rangle - \langle \Delta x(\Delta t) \rangle^2 + V_T^2 \Delta t^2 \quad . \tag{11}$$

We expect the new squared term $V_T^2 \Delta t^2$ to be carried out to the total MSD equation for $\Delta \vec{r}$ (we do not demonstrate this step because it would require the full solution for the adapted equation system 3-5):

$$MSD' = MSD + V_T^2 \Delta t^2 \quad . \tag{12}$$

We can now import the MSD result from Eq. 6, yielding

$$\langle |\Delta \vec{r}|^2 \rangle = 2D(\Delta t - P(1 - e^{-\Delta t/P})) + \frac{2DS}{1-S}\Delta t + B\Delta t^2 \quad , \tag{13}$$

where $\sqrt{B} = V_T$. This justifies the use of an extra ballistic term in our MSD fitting curve, which does fit the simulation results as you will see in the paper's result section.

At last, it is important to state that robust characterization is key in biomedicine. For example, when studying cancer cell aggressiveness and drug interventions, metric's robustness determines the quality of results, which in turn determines the reliability of medical reports and interventions. For this purpose, mathematical models provide parameters with clear interpretation and measurement. In the case of chemotaxis, the parameter "terminal velocity" can be used to quantitatively compare two experiments and decide if a drug hindered or intensified a cell's response to a chemical field.

C. Cells in a Lattice: the CompuCell3D Implementation of Cellular Potts Model

A development from Ising and Potts models of Solid State Physics, Cellular Potts Model (CPM) allows a set of pixels to behave as an object with size and shape, interacting with other objects around [1–3]. Originally, CPM was not specifically designed for use in the field of Cellular Biology. However, it has been found to be unexpectedly useful in this area, and it is now one of the primary applications of the model. Despite this, CPM still has other applications in fields such as materials science.

In Ising model [24], originally intended for ferromagnetism, each pixel of the lattice can transit between two states (spin up and down) depending on its neighbors. Two neighboring pixels of same spin have a lower energy than two neighboring pixels of opposite spins. Let's propose an algorithm that picks a random pixel in the grid and offers a state switch. If global energy decreases, the state switch is accepted. Otherwise, a Boltzmann Probability

$$P_{Boltzmann} \propto e^{-\Delta E/T} \tag{14}$$

will decide if it switches or not based on energy variation and temperature. The higher the temperature, the higher the probability to accept a transition that would increase global energy. If we repeat this process over many steps, the system will evolve pursuing an energy minimum. As we set higher temperatures, the system can occupy higher energy states, increasing entropy.

In Potts model [25], the algorithm is similar to Ising's, but more than 2 states are possible. A simple use of Potts model is to simulate infectious disease spread in a lattice, where each pixel represents a house. Consider three pixel states: healthy, infected and removed. The evolution can be implemented similarly to Ising model, the difference being that each state switch can opt between more than one transition, for example: an infected site could recover without immunity (susceptible), or could become permanently immune or die (removed).

CPM [1] distinguishes itself from its predecessor by implementing Metropolis-Hastings algorithm [26] differently; pixels alone can not switch state (no matter the temperature), instead, one pixel attempts to copy its attributes to a neighbor. In other words, if a pixel of state A is surrounded by pixels of state A, it will never switch to B even for very high temperatures, because there is no neighbor in state B to attempt a copy. The effect is to

favor connected clusters with pixels of the same state. State switches will happen only in the interface between neighboring clusters. In what follows, we call "cell" each cluster with pixels of the same state, reserving the term "cluster" when more than one cell compose a bigger structure.

To give an example of a simple CPM simulation, consider a 2D square lattice where different cells are located. The evolution of the system goes as follows:

1. A random pixel in the $N \times N$ 2D square lattice is chosen. Let $\vec{i} = (x, y)$ be its position.

2. A random pixel is picked in the neighborhood of pixel $\vec{i}$. Let $\vec{j}$ be the position of the randomly picked neighbor pixel.

   It is possible to define the neighbor order. In general, people use neighbor order equal to 2, so the neighboring region of a pixel is the 8 closest neighboring pixels (in 2D). The central pixel is not counted as part of the neighborhood.

3. Get the state σ of pixel $\vec{i}$, where σ stands for the cell to which pixel $\vec{i}$ belongs.

   In cases with different cell types, another attribute, τ, will represent the cell type, and the pixel state will become the list of attributes (σ, τ). Other attributes can be defined for more complex simulations.

4. Calculate the global energy in the current configuration and in the configuration supposing pixel $\vec{j}$ acquires all of $\vec{i}$'s attributes (complete overwrite). There are two options:

   $\Delta E_{total} \leq 0$) If global energy decreases, pixel $\vec{j}$ switches state and become identical to pixel $\vec{i}$, acquiring all of $\vec{i}$'s attributes.

   $\Delta E_{total} > 0$) If global energy increases, state switch can be accepted with Boltzmann probability (Eq. 14);

5. The above process is repeated N_{MCS} times. $N_{MCS} = N \times N$ defines a Monte Carlo Step (MCS), which means the amount of copy attempts is equal to the number of lattice pixels. The MCS is the time unit in a CPM simulation.

The global energy equation is the main determinant of the system's behavior. One of the simplest energy equations in CPM considers a volume energy (to constraint cell sizes) and

a contact interaction energy term (analogous to surface tension in liquids):

$$E_{total} = \sum_{\vec{i}} \sum_{\vec{j}} J(\tau_{\vec{i}}, \tau_{\vec{j}})(1 - \delta(\sigma_{\vec{i}}, \sigma_{\vec{j}})) + \sum_{\sigma} \lambda_{\sigma} (V_{current}^{\sigma} - V_{target}^{\sigma})^2 \quad . \qquad (15)$$

Again, σ is the cell identifier, and τ is the cell type identifier. J is the energy per edge between neighboring pixels. J depends on the types of each pair of pixels sharing a surface, thus it is a symmetrical matrix listing each possible interaction between all pairs of cell types in a simulation. Each cell has a target volume V_{target}^{σ}, so they can grow and shrink individually. The parameter λ is the inverse of compressibility, it regulates how much the cell volume $V_{current}$ can fluctuate around V_{target}. Now we will present the softwere CompuCell3D (CC3D) and use it to build a simple model of Cellular Biology as an example of the above CPM energy equation.

CompuCell3D (CC3D) (accessible at compucell3d.org) [27] is an open software that runs a CPM algorithm (in C++) as a black box, but allows the user to intervene between each MCS (in Python) and initialize parameters and other useful simulation Plugins (in XML and Python). Users can add precoded energy terms to the global energy equation. It is possible to implement: cell attraction, cell surface, force over cells, cell interactions with fields and others. CC3D offers: 1) diffusion solvers for chemical fields, giving cells the ability to secrete, uptake and diffuse; 2) methods to model mitosis and cell death; 3) and Plugins to manipulate data acquisition, data output, parameter scan etc. CC3D has a Player with real time visualization display of cells and fields in the lattice, and the user can plot simulation data in real time in the Player. The code editor Twedit++ provides examples on how to implement various useful methods and helps managing all simulation files. Twedit++ is also included in the CC3D package. Nowadays, CC3D can be called as a Python library, giving complete freedom for modellers to use it inside Python scripts.

CC3D headquarters is placed in Indiana University, Bloomington, USA. The multicultural team of instructors and developers is led by Prof. James A. Glazier (Indiana University, Bloomington), and includes members and ex members from our research group LabCel (Instituto de Física, Universidade Federal do Rio Grande do Sul). Prof. Gilberto L. Thomas (LabCel, IF-UFRGS) leads the collaboration between LabCel and CC3D. CC3D team holds yearly virtual Workshops open for any graduate and undergraduate student of any science backgrounds from any country and university. Students learn basic concepts of modelling,

basic Python, and basic to advanced use of CC3D software with examples and practice. At the end of the course, students form groups and apply learnt content in problems of real importance to scientific community. Visit CC3D site (compucell3d.org) to learn more.

Now we will show how to build a simple self segregation model step-by-step in CC3D. If you already have CC3D installed (any version newer than 4.0.0), you should be able to follow the steps and see the results. But first, it is important to understand what we are trying to model and what we want to achieve. Self segregation is in our day-by-day life: oil and water is a common example, football fans in a stadium is another one, thank god, and cornstarch and milk is probably the most annoying of all. In Cellular Biology, cell segregation is complex because it can result from very different processes: cell-cell communication via chemical fields, difference in adhesion between different pairs of cell types, cell tolerance to external medium, difference in motility between different cell types, and even cell phenotype differentiation due to diffusive fields or contact with surrounding agents [28].

With the energy equation 15 we are limited to study difference in adhesion and tolerance to external medium, however, it is possible to implement and simulate all the enumerated processes in CC3D. In our example, starting with two different types of cells, our goal is to show which conditions are enough to promote segregation and to observe different manifestations of segregation depending on the simulation parameters.

1. Go to CC3D installation folder and open Twedit++;

2. Open the tab menu "CC3D Project";

3. Click "New CC3D Project...";

4. Give your simulation the name "CellSeg", keep "Python+XML" option checked in "Simulation Type", and hit "Next";

5. You can change the lattice size, boundary conditions (for cells, not fields), lattice geometry (square or hexagonal), Average Membrane Fluctuation (this is the Potts temperature that appears in Eq. 14), the Neighbor Order of pixel copies, and the total number of MCSs (simulation length). For now, we will only change the Neighbor Order to 2, leaving the rest as it is;

6. Check option "Blob" in the Initial Cell Layout, so that cells are initialized in a spherical blob instead of a square, hit "Next";

7. In the "Cell Type" text bar, type "CellA" and hit "Add". Then type "CellB" and hit "Add". Now we have set our two cell types CellA and CellB. Cell type "Medium" is a default cell type which represents the surrounding cell medium. DO NOT REMOVE MEDIUM TYPE. Hit "Next";

8. Leave "Chemical Fields" section as it is. Hit "Next";

9. In "Cell Properties and Behaviors" section, check boxes "Contact" and "VolumeFlex". Leave the rest as it is. Hit "Next";

10. Hit "Finish".

We have our simulation set. Now let's run it and see what we got.

1. In Twedit++, in the "CC3D Project" sub window, double click our project "CellSeg.cc3d". Twedit++ will open 3 tabs with our main code files.

   "CellSeg.xml" is our setup file. In "CellSegSteppables.py", we are free to code whatever we want in the starting simulation step (start function), between MCSs (step function) and at the end (finish function). At last, "CellSeg.py" is the main python file that calls each class inside "CellSegSteppables.py". We are not changing these files for this example.;

2. In Twedit++, in the "CC3D Project" sub window, right click our project "CellSeg.cc3d";

3. Click "Open in Player". A new window will open with the cell visualization display and other options.

   You can hit pause, stop and run at the top left corner. By default, when you click "Open in Player", the simulation automatically starts running. For now, we only get our cell initialization and random fluctuations at the cell borders. It is time to play with it;

4. In the "Model Editor" sub window, open the drop down arrow of "Plugin Contact". Then open all drop downs inside it. The default value 10 showing in each pair of cell types is our energy term $J(\tau_1, \tau_2)$;

   All energies are equal, CellA and CellB have the same behavior and will avoid contact with Medium, keeping the cluster tight. The Medium avoidance comes from

the fact that only one interface between CellA and CellB is energetically more favorable than having a strip of Medium between CellA and CellB, which has one additional interface with same contact energy value;

5. Leave the simulation running and set $J(CellA, CellA) = 1.0$ and $J(CellB, CellB) = 1.0$: double click the value 10.0, change the value to 1.0 (IT HAS TO BE A FLOAT) and press enter.

   With less contact energy between cells of same type, it becomes more favorable to establish clusters of cells of same type. This is our first manifestation of self segregation by different cell adhesion. Some small clusters of cells will remain alone because there is no long range interaction. However, for very long simulation times, the tendency is that only a few clusters with many cells remain;

6. Set $J(CellA, CellA) = 10.0$ and $J(CellB, CellB) = 10.0$ again (IT HAS TO BE A FLOAT) and see the cell clusters relaxing due to reduction in surface tension;

7. Now set $J(Medium, CellA) = 100.0$ (IT HAS TO BE A FLOAT), this will make CellA intolerant to Medium, so cells of this type will avoid contact with Medium.

   Very quickly CellB surrounds CellA, producing another manifestation of self segregation. This one, in particular, is important for embryo development, as embryo needs to segregate endoderm and ectoderm.

8. Have fun. You can test different combinations of contact energies to produce other types of segregation.

   We did not code directly in the python files and already got some interesting results. If we were to use the Python files to their full potential, we could implement cell communication through chemical fields, cell oxygen depletion inside the aggregate, cell differentiation via contact or via chemical fields, cell growth and division, cell death, cell migration, chemical reaction pathways, and many other phenomena.

This simple model is our preferred example of an application of CPM to cellular biology. CPM cells' fluctuations and damped medium have great resemblance with experimental observations, justifying its wide use for Cellular Biology rather than materials science. More-
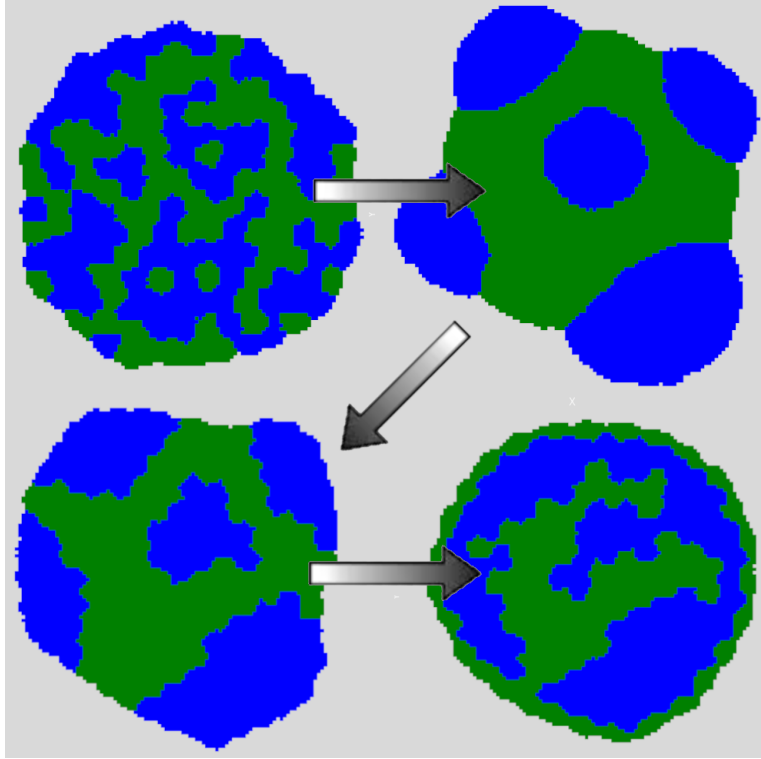
Figure 5. This is the transition from the cell initialization state (first screenshot), passing through the first segregated state with CellA and CellB grouping together (step 5), back to relaxed state (step 6), and finally to the engulfment state where CellB surrounds CellA (step 7).

over, cell behaviors like contact and volume energies alone achieve impressive results given their simplicity. If you did not follow the exercise, take a look at the results in Fig. 5.

This ends the theoretical introduction. Now you are able to read and fully understand our paper "CompuCell3D Model of Cell Migration Reproduces Chemotaxis".

CLOSURE

Our simulation presented in the paper reproduced chemotactic behavior from a reorientation dynamics exclusively. We demonstrated the presence of a drift speed (net positive velocity in the polarization direction), a required feature for this mechanism to work. We detected diffusion in cell velocity due to membrane fluctuations, driven by CPM stochastic evolution. Finally, we characterized all notable cell behaviors quantitatively from position and polarization data. We highlight:

1. The use of a modified Fürth equation for MSD (Mean Square Displacement), with a

terminal speed term that also appears in mVACF (Mean Velocity Auto Correlation Function);

2. The definition of four time intervals that mark transitions between different MSD regimes;

3. The analysis of cell displacement distributions in both laboratory and cell reference frames;

4. The ratio between terminal speed and drift speed defining chemotactic efficiency.

All these metrics can be applied to experiments (provided that polarization is measured at all trajectory steps) and they allow quantitative comparison between data from different sources.

The most important concept in this paper is polarization. I take the risk to state that any cell migration model must include polarization to be realistic, and any cell migration analysis must include polarization to be complete. In our CPM model, we create Lamel (lamellipodium) pixels in the cell to promote polarization reorientation. Other modelling methods (center, vertex, phase field) can easily implement analogous dynamics. Mathematical models are not that simple. Traditional Langevin-like equations are insufficient to model realistic chemotaxis. Developing robust mathematical models of chemotaxis requires the use of 2-dimensional stochastic differential equations that explicitly account for the cell's polarization direction, incorporate velocity drift in the direction of polarization, and model chemotaxis as a factor that influences polarization orientation, rather than pulling the cell up gradient.

In immune response and cancer spread simulations, appropriate chemotaxis dynamics can significantly impact the quality of results and conclusions. Hence, our next step could be to show the differences between our version and the traditional version of chemotaxis in these types of problems. Gilberto Lima Thomas, together with new students, is developing a multicellular simulation, paving the way to the study of wound healing or other collective migration problems. In the scope of theoretical models, our results help Rita M. de Almeida and Guilherme Y. Shoiti in their mathematical model of cell migration, still a work in development.

ACKNOWLEDGMENTS

BIBLIOGRAPHY

[1] François Graner and James A. Glazier. Simulation of biological cell sorting using a two-dimensional extended potts model. *Physical Review Letters*, 69:2013–2016, 1992.

[2] James A. Glazier and François Graner. Simulation of the differential adhesion driven rearrangement of biological cells. *Physical Review E*, 47:2128–2154, 1993.

[3] Paulien Hogeweg. Computing and organism: on the interface between informatic and dynamic proceses. *Biosystems*, 64:97–109, 2002.

[4] S.R. Bolsover, J.S. Hyams, E.A. Shephard, H.A. White, and C.G. Wiedemann. *Cell Biology: A Short Course*. Wiley, 2004.

[5] Douglas A Lauffenburger and Alan F Horwitz. Cell migration: A physically integrated molecular process. *Cell*, 84(3):359–369, feb 1996.

[6] A. M. Turing. The chemical basis of morphogenesis. *Phil. Trans. R. Soc. Lond. B*, 237:37–72, 1952.

[7] Kristen F. Swaney, Chuan-Hsiang Huang, and Peter N. Devreotes. Eukaryotic chemotaxis: A network of signaling pathways controls motility, directional sensing, and polarity. *Annual Review of Biophysics*, 39(1):265–289, apr 2010.

[8] Orion D. Weiner, Paul O. Neilsen, Glenn D. Prestwich, Marc W. Kirschner, Lewis C. Cantley,

and Henry R. Bourne. A PtdInsP3- and rho GTPase-mediated positive feedback loop regulates neutrophil polarity. *Nature Cell Biology*, 4(7):509–513, jun 2002.

[9] Chang Y. Chung and Richard A. Firtel. Signaling pathways at the leading edge of chemotaxing cells. *Journal of Muscle Research and Cell Motility*, 23(7/8):773–779, 2002.

[10] Carol L. Manahan, Pablo A. Iglesias, Yu Long, and Peter N. Devreotes. Chemoattractant signaling in dictyostelium discoideum. *Annual Review of Cell and Developmental Biology*, 20(1):223–253, nov 2004.

[11] Tatyana M. Svitkina and Gary G. Borisy. Arp2/3 complex and actin depolymerizing factor/cofilin in dendritic organization and treadmilling of actin filament array in lamellipodia. *Journal of Cell Biology*, 145(5):1009–1026, may 1999.

[12] Miguel Vicente-Manzanares and Alan Rick Horwitz. Cell migration: An overview. In *Methods in Molecular Biology*, pages 1–24. Humana Press, 2011.

[13] Isabela C. Fortunato and Raimon Sunyer. The forces behind directed cell migration. *Biophysica*, 2(4):548–563, dec 2022.

[14] A Mogilner and Kinneret Keren. The shape of motile cells. *Current Biology*, 19:R762–R771, 2009.

[15] Evanthia T. Roussos, John S. Condeelis, and Antonia Patsialou. Chemotaxis in cancer. *Nature Reviews Cancer*, 11(8):573–587, jul 2011.

[16] Tian Jin, Xuehua Xu, and Dale Hereld. Chemotaxis, chemokine receptors and human disease. *Cytokine*, 44(1):1–8, oct 2008.

[17] Henry R. Bourne and Orion Weiner. Cell polarity: A chemical compass. *Nature*, 419(6902):21–21, sep 2002.

[18] G. Wayne Brodland. How computational models can help unlock biological systems. *Seminars in Cell and Developmental Biology*, 47-48:62–73, dec 2015.

[19] John Von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton, NJ, USA: Princeton University Press, 1944.

[20] Emanuel F. Teixeira, Heitor C. M. Fernandes, and Leonardo G. Brunnet. A single active ring model with velocity self-alignment. *Soft Matter*, 17:5991–6000, 2021.

[21] Gilberto L Thomas, Ismael Fortuna, Gabriel C Perrone, Glazier James A, Julio M Belmonte, and Rita M C de Almeida. Parameterizing cell movement when the instantaneous cell migration velocity is ill-defined. *Physica A*, 550:124493, 2020.

[22] Rita M C de Almeida, Guilherme S Giardini, Mendeli Vainstein, James A Glazier, and Gilberto L Thomas. Exact solution for an anisotropic ornstein-uhlenbeck process. *http://arxiv.org/abs/2002.00051*, 2020.

[23] Włodzimierz Korohoda, Zbigniew Madeja, and Jolanta Sroka. Diverse chemotactic responses of *Dictyostelium discoideum* amoebae in the developing (temporal) and stationary(spatial) concentration gradients of folicacid, camp, Ca^{2+} and Mg^{2+}. *Cell Motil Cytoskeleton*, 53:1–25, 2002.

[24] A.B. Bortz, M.H. Kalos, and J.L. Lebowitz. A new algorithm for monte carlo simulation of ising spin systems. *Journal of Computational Physics*, 17(1):10–18, 1975.

[25] F. Y. Wu. The potts model. *Rev. Mod. Phys.*, 54:235–268, Jan 1982.

[26] W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, apr 1970.

[27] Maciej Swat, Gilberto L. Thomas, Julio Monti Belmonte, Abbas Shirinifard, Dimitrij Hmeljak, and James A. Glazier. Multi-scale modeling of tissues using compucell3d. *Methods in Cell Biology*, 110:325–366, 2012.

[28] S. F.Gabby Krens and Carl-Philipp Heisenberg. Cell sorting in development. In *Current Topics in Developmental Biology*, pages 189–213. Elsevier, 2011.