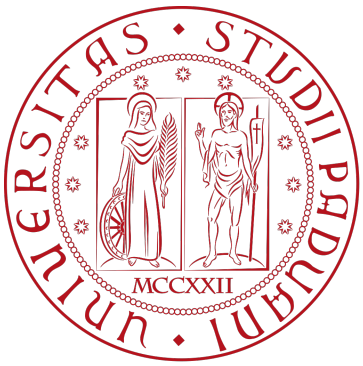


UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Master Degree in Physics

Final Dissertation

A stochastic modeling for the emergence of loners in social amoebae

Thesis supervisor

Prof. Antonio Trovato

Candidate

Antonio Santo

Academic Year 2021/2022

Contents

1 Introduction

In the following text I will discuss a quorum-sensing model recently proposed, which has the purpose of describing the aggregation mechanism involved in the life cycle of the cellular slime mold *Dictyostelium discoideum*. In particular, the object of interest of such model and my work are the individuals left out of this synchronization, called loners. After a brief introduction of the model and the type of data that can be produced in computer simulations starting from it, I will underline some of its aspects that can be upgraded and propose solutions for some of them. Then the results of such implementations will be presented, alongside some trials regarding the problems of the model which instead remained unsolved. Finally, in the last section I will present an analysis of the data made by using an empirical function for the time evolution of the population of the amoebae. At the end I will propose a way to help the identification of the molecule responsible for the quorum-sensing mechanism.

1.1 Loners

All the living organisms with which we are most familiar start their life cycle as a single cell, usually a fertilized egg, which begins to grow through repeated cell divisions, until achieving a multicellular state. During this development they don't have the means to take in energy by themselves, but receive nutrients from an external source like a yolk. By the time this stock of energy is depleted, such organisms are able of feeding on their own.

Cellular slime molds, like the well studied *Dictyostelium discoideum*, instead behave very differently. They begin their life cycle emerging from spores as unicellular amoebae, which immediately start to feed [2]. As they eat they will also begin to divide, growing in number. Once they have cleaned the area of all the available food, the prolonged starvation triggers an aggregation mechanism, in which the amoebae stream together to become multicellular. When this mound is

complete, the slime mold enters a slug phase, in which the complex of amoebae migrates from the feeding site to an area ideal for the formation of fruiting bodies. In this stage they can no longer do anything that requires an intake of energy, and they just wait for the spores to be transported away by some passing by animal, starting the process from the beginning.

One of the most striking and interesting aspects of the life cycle briefly described above is of course the coordination manifested by the amoebae in their aggregation. Such type of collective behavior is not unique to the slime molds, but can be found frequently in nature. And to be common it is also the presence of imperfections in these synchronization mechanisms, in other words "out-of-sync" individuals. These loners were often seen just as byproducts of large-scale coordination attempts and only sometimes considered as a possible trait capable of being shaped by evolution.

Recently it has be shown [1] that, in the case of *Dictyostelium discoideum*, loners are indeed an heritable feature and that it is present variability in thier expression among different genetic variants of the same species. In fact it was observed, in repeated experiments in which different strains of *D. discoideum* were involved, that loners densities produced by a given strain fell consistently within a conserved distribution. In particular the mean and variance of such distribution were significantly different with respect to those relative to the loners distributions of other strains. It is known that loners persist temporarily after the aggregation process ends, and that if they are able to find a food source they will eat and divide, re-enacting the multicellular develompment. For this reason it would be possible for evolutionary processes to act on them.

But the findings presented in [1] answer also to a very important question. While for low initial densities of cells the amoebae were too sparse for aggregation to occur (hence making the whole population correspond to loners), above a threshold aggregation occurred with increasing efficiency, and loners densities decreased. At high initial cell densities the loners densities plateaued, indicating that it is the number (or density) of loners which is heritable. But this means that the

"decision-making" process underlying the aggregation mechanism is not just a context-independent stochastic switch, otherwise it would have been the fraction and not the number of loners to be heritable. What these observations suggest is instead the presence of a quorum-sensing process. In order to describe and investigate the properties of the aggregator-loner partioning, Tarnita *et al.* proposed a spatially explicit individual-based model which is presented below.

1.2 The model

The model is built around the central hyphotesis that the imperfect synchronization of the population of amoebae comes from the stochasticity of the quorum sensing mechanism. A square system of lateral length ℓ is considered, which represents an aggregation territory, while time is discretized with timestep $\mathrm{d}t$. The simulation is started with N_0 cells placed randomly with a uniform distribution, assumed to be in the preaggregating state P. Such P-cells do not move and emit signaling molecules at a strain-specific rate γ.

The signal diffuses with diffusion coefficient D and regulates the response of individual cells to the surrounding population: if a cell senses a signal density above a strain-specific sensitivity threshold θ, then it has a probability rate λ, this too strain-specific, of transitioning to the aggregating state A. The A-cells move towards the center of the square box (which is the exogenously imposed aggregation site) with a constant strain-specific velocity v and emitting signal molecules at the same rate γ of the P-cells. When they arrive to the aggregation site, the A-cells adhere to the mound transitioning to M-cells. In this multicellular stage the amoebae do not move and stop emitting signal, ceasing also to sense it like the A-cells.

Simulations are allowed to run until the time between two consecutive arrivals to the mound is larger than a fixed parameter t_{arr}. The final density of loners is computed as an average over

different trials of the number of remaining P-cells for each realization, divided by the area of the system. The signal density at a time t and position $\mathbf{r}$ of a focal P-cell is obtained by solving the generalized diffusion equation:

$$\sigma(\mathbf{r},t) = \sum_{\mathbf{r}'\neq\mathbf{r}} \sigma_{\mathbf{r}'}(|\mathbf{r}-\mathbf{r}'(t)|) \tag{1.1}$$

Here $\sigma_{\mathbf{r}'}$ indicates the individual contribution of a cell at position $\mathbf{r}'$, and the sum runs over the locations of the other P-cells and A-cells. In particular the presence of amoebae in the A-state is what gives the time dependance to $\mathbf{r}'(t)$.

In order to determine $\sigma_{\mathbf{r}'}$, the assumptions made regarding the signal are: that it is continuosly released by each cell with rate γ, that diffuses with diffusion constant D, and that it spontaneously decays at a degradation rate η. Also, for the purpose of mimicking the presence of multiple aggregation territories and hence diffusion of signal from outside our square box, we impose periodic boundary conditions. Considering the cells as punctual sources, the spatiotemporal evolution of the signal's density emitted from a cell at the origin is obtained by solving the generalized diffusion equation

$$\frac{\partial\sigma_{\mathbf{0}}}{\partial t}(x,y,t) = D\nabla\sigma_{\mathbf{0}}(x,y,t) - \eta\sigma_{\mathbf{0}}(x,y,t) \tag{1.2}$$

But for simplicity we consider first a system of infinite size and then embed the solution in a periodic space. In order to solve the differential equation we impose as boundary conditions the continuos emission of signal and that the signal's density goes to zero as the distance from the emitting cells tends to infinity. Since this makes our problem radially symmetric we can use polar coordinates and consider the stationary limit:

$$D\left(\frac{\mathrm{d}^2\sigma_{\mathbf{0}}}{\mathrm{d}r^2}(r) + \frac{1}{r}\frac{\mathrm{d}\sigma_{\mathbf{0}}}{\mathrm{d}r}(r)\right) - \eta\sigma_{\mathbf{0}}(r) = 0 \tag{1.3}$$

Being $r = 0$, the position of the emitter, a singular point, we first assume the cells to have a finite radius $\tilde{r}$ and then, after finding the solution, take the $\tilde{r} \to 0$ limit. In this way the boundary conditions can be written as

$$\begin{cases} \sigma_{\mathbf{0}}(r \to \infty) = 0 \\ -2\pi\tilde{r}D\frac{\mathrm{d}\sigma_{\mathbf{0}}}{\mathrm{d}r}(r)\big|_{r=\tilde{r}} = \gamma \end{cases} \tag{1.4}$$

Now Equation (1.3) is the modified Bessel equation of order zero, which has as general solution

$$\sigma_{\mathbf{0}}(r) = A\,\mathrm{I}_0(\sqrt{\frac{\eta}{D}}r) + B\,\mathrm{K}_0(\sqrt{\frac{\eta}{D}}r) \tag{1.5}$$

where I_0 and K_0 are the zero order modified Bessel functions respectively of the first and second kind. The argument of both these functions contains the characteristic length $\sqrt{D/\eta}$, which quantifies the interplay between diffusion and degradation. In fact it can be interpreted as the typical distance over which signal molecules can diffuse before being degraded.

Imposing the two boundary conditions it's possible to obtain

$$\begin{cases} A = 0 \\ B = \frac{\gamma}{2\pi\tilde{r}\sqrt{\eta D}\mathrm{K}_1(\sqrt{\frac{\eta}{D}}\tilde{r})} \end{cases} \tag{1.6}$$

In fact $\mathrm{I}_0(r)$ diverges as $r \to \infty$; while B was obtained using the fact that $\mathrm{K}_0{}'(r) = -\mathrm{K}_1(r)$, where $\mathrm{K}_1(r)$ is the first order modified Bessel function of the second kind. Thus putting now all together and taking the $\tilde{r} \to 0$ limit finally gives:

$$\sigma_{\mathbf{0}}(r) = \frac{\gamma}{2\pi D}\,\mathrm{K}_0(\sqrt{\frac{\eta}{D}}r) \tag{1.7}$$

As previously said this solution assumes infinite system size, but can be easily extended to a finite domain with periodic boundary conditions, which means working on an infinite lattice made

of repeated copies of the focal tiles. The signal's density within such focal tile is obtained adding over the contributions of all the infinite mirrored tiles. But since the parameters used are such that the characteristic length $\sqrt{D/\eta} << \ell$, and since $\mathrm{K}_0(r) \to 0$ as $r \to \infty$, it follows that $\sigma(\ell/2) \approx 0$ and we can truncate the sum over tiles at the nearest neighbors of the focal one. This is equivalent to calculating the distance $r = \sqrt{r_x^2 + r_y^2}$, between the focal cell and an emitting cell (x_{em}, y_{em}), as:

$$r_x = \begin{cases} |x - x_{em}| & \text{if } |x - x_{em}| \leq \ell/2 \\ \ell - |x - x_{em}| & \text{if } |x - x_{em}| > \ell/2 \end{cases} \tag{1.8}$$

$$r_y = \begin{cases} |y - y_{em}| & \text{if } |y - y_{em}| \leq \ell/2 \\ \ell - |y - y_{em}| & \text{if } |y - y_{em}| > \ell/2 \end{cases} \tag{1.9}$$

2 Discussion about the model

2.1 The simulations

In order to produce simulations similar to the ones showed in the supplementary material **S5 Fig.** in [1], I wrote a *Fortran 90* code. The program creates a uniform distribution of amoebae in a square box of length $\ell = 0.2$ cm, then stores the signal density exchanged by each pair of cells in a lower triangular matrix $N_0 \times N_0$. Indeed what matters is the local concentration of signal seen by a single P-cell. Since the density of signal emitted by each amoeba is a radial function, I can treat it as a 2-body interaction.

Each element of my matrix is computed using periodic boundary conditions and the `BESK0` routine for the modified Bessel function, written by John Burkardt. Then for each P-cell, once the local density of signal is obtained, the code checks if the threshold θ is surpassed or not. In case of positive response a random number between 0 and 1 is drawn, if it is lower than $\lambda \mathrm{d}t$ a P-to-A transition occurs. Eventually, as the simulation time advances with timestep $\mathrm{d}t = 0.01$ hours, the amoebae start travelling towards the center of the tile with velocity v. When they reach the inside of the disc of radius $v\mathrm{d}t$ centered in the aggregation point, the A-to-M transition takes place.

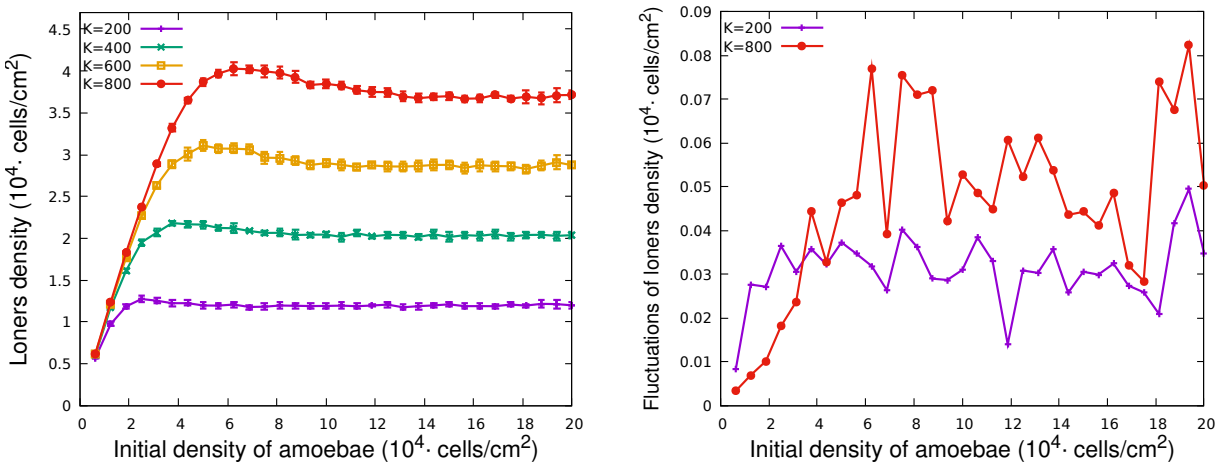
This goes on as long as the population of P- and A-cells is able to mantain a local density of signal above the threshold value θ. Then if the time between two consecutive arrival to the mound becomes larger than $t_{arr} = 1$ hour, the simulation is stopped. The process is repeated using the same number of initial cells N_0 in order to assign to each N_0 an average number of loners over different realizations and a standard deviation. More precisely I averaged over 10 trials instead of 100, as done in [1], because it would have been too time consuming. For all the parameters involved in the simulations, as the diffusion coefficient D or the degradation constant η, the values provided in the supplementary material of [1] were taken, and they are reported in Table 1.

The choice of which value to assign to each quantity is important also for a very specific reason. The model proposed in the article contemplates the possibility of having different populations of *Dictyostelium discoideum* that behave differently in the same external conditions. Such variety is reached through a series of parameters that can be varied in order to characterize a strain. In particular these effects can be condensed in two ratios: λ/v and θ/γ. Indeed what affects the total loners density is for sure how quickly P-cells switch to the A-state with respect to how fast they are depleted of signal. The larger is the velocity v at which the amoebae travel toward the aggregation point, the narrower is the time window in which the quorum is mantained, resulting in more loners left behind. Conversely a larger transition rate λ has the effects of synchronizing the switches to the A-state, so that in the same time interval I have more P-to-A transitions with respect to a lower λ, and thus fewer loners. The result of this interplay is that the farther a cell is from the aggregation center, the sooner it is left without a quorum, and the more likely it is to become a loner. At the same time the ratio θ/γ, between the quorum threshold and the signal's production rate, represents how easy it is to mantain a quorum. Given two strains in two environments with the same properties (i.e. same D and η), the one which produces more signaling molecules will result being a better aggregator, which means leaving fewer loners behind. Instead having an higher threshold will make it easier for the P-cells population not to be activated and thus to be left without a quorum.

During my tests I choose to vary only the θ/γ ratio, which will be referred to as K. This because in principle changing the two ratios should be qualitatively equivalent, and also because the condition for ending the simulations is inevitably susceptible to the value of λ/v. Since the

	D	η	λ	**v**	$\mathbf{t_{arr}}$	ℓ	**dt**
Units	cm^2/hours	hours^{-1}	hours^{-1}	cm/hours	hours	cm	hours
Value	$3.6 \cdot 10^{-4}$	72	1	$7.2 \cdot 10^{-2}$	1	0.2	0.01

Table 1: *Table containing the parameters used in the simulations.*



(a) *Final loners density as function of the initial density of plated cells. The error bars correspond to the standard deviation over 10 trials.*

(b) *Standard deviations over 10 trials of the loners density. Only the curves for two values of K are reported for clarity reasons.*

Figure 1: *Data obtained from the* Fortran 90 *code. The values used for the various parameters are the same given in the supplementary material of [1].*

modification made to the model, that will be discussed in the next subsection, shed some doubts on the t_{arr} criterion, I felt it was more appropiate to focus just on K.

In the end I obtained curves very similar to the ones produced in the article. After an initial linear increase for low initial densities, signifying the quorum isn't met, the loners density starts to decrease until it stabilzes at a constant value for high densities of plated cells. The value of this plateau is strain-specific, as showed in Figure 1a, and for simplicity it is considered to begin, as a convention, for initial densities $\geq 12.5 \cdot 10^4$ cells/cm^2. This last point, while had the consequence of making me neglect points already inside the plateau zone for lower values of K, it also allowed me to automatize a lot of the computations that I will be presenting.

Always referring to Figure 1a, here I used the standard deviation as error for the final loners density corresponding to each density N_0/ℓ^2 of initially plated cells. Such choice is dictated by the fact that I'm simply interested in having an index of the variability of the loners distribution for

each strain.

Once that it has been confirmed that my code reproduced well the model, some upgrades to the simulations were introduced.

2.2 The limits of the model

The model proposed by Tarnita *et al.* is of course a heavy semplification of the complicated processes lying behind the aggregation mechanism of *Dictyostelium discoideum*, and as such it has several limits. A very important one is that it is unable to reproduce the strain specific statistics, as noted also in [1]. In fact from the experimental data it is possible to see that different strains have not only a different value of the plateau but also very different values of the relative fluctuations around it. While it is true that in the simulations the standard deviation, as the plateau, grows with K (Figure 1b), such growth is very slow compared to the one observed experimentally.

2.2.1 Stationary condition

Another point concerns the reaching of stationarity: the evolution of the signal's density is described by a diffusion equation with a degradation term η. Typically in these cases the steady state is reached with a timescale given by the inverse of the degradation coefficient, but the parameters used in [1] are such that $1/\eta > \mathrm{d}t$. One could think of implementing the simulations using the time-dependant solution and then confront the results. It is possible to find such solution and it has the form [1]:

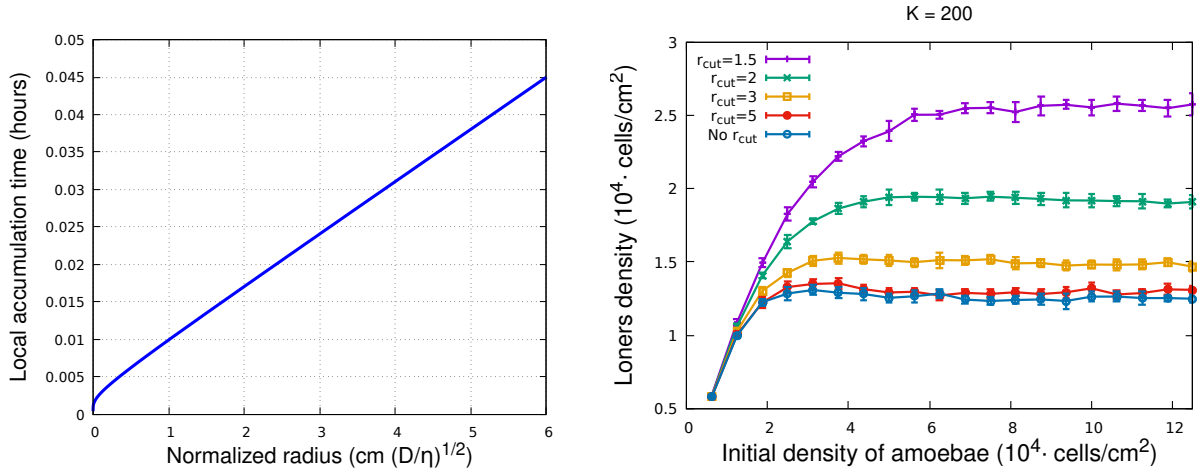
$$\sigma_{\mathbf{0}}(r,t) = \frac{\gamma}{2\pi D}\Big[\mathrm{K}_0(\sqrt{\frac{\eta}{D}}r) - \frac{e^{-\eta t}e^{-r^2/4Dt}}{2Dt}\,Z(r,t)\Big] \tag{2.1}$$

with

$$Z(r,t) = \int_0^\infty \mathrm{d}r_0\, r_0\, \mathrm{K}_0(\sqrt{\frac{\eta}{D}}r_0)\mathrm{I}_0(\frac{rr_0}{2Dt})e^{-r_0^2/4Dt}. \tag{2.2}$$

[1] The complete calculations with which this expression was obtained are reported in Appendix A.

Unfortunately simulations made with this expression for the signal's density would be too time consuming because of the presence of the integral involving the modified Bessel functions. If one wish to keep the stationary solution it is necessary, for example, to take a larger timestep.



(a) *Local accumulation time relative to (1.2) as a function of the distance from the source.*

(b) *Loners density for $K = 200$ and using different cutoff radii. Each value of r_{cut} is in $\sqrt{D/\eta}$ cm units.*

Figure 2: Plots relative to the discussion about the stationary solution.

A useful tool for this purpose is the concept of local accumulation time [4]. Indicating as $\sigma^s_{\mathbf{0}}(r)$ the steady state signal density, it's possible to define

$$R(r,t) = \frac{\sigma^s_{\mathbf{0}}(r) - \sigma_{\mathbf{0}}(r,t)}{\sigma^s_{\mathbf{0}}(r)}, \tag{2.3}$$

which is called the local relaxation function and describes the approach of $\sigma_{\mathbf{0}}(r,t)$ to $\sigma^s_{\mathbf{0}}(r)$ at a certain distance r from the source. For any distance $r > 0$ it starts from unity and tends to zero as the time passes. In this way considering two different times $t_1 < t_2$, the difference $R(r,t_1) - R(r,t_2)$ can be viewed as the fraction of the steady state level that has been accumulated between t_1 and t_2. It follows that $-(\partial R(r,t)/\partial t) \cdot \mathrm{d}t$ corresponds to the fraction of steady state accumulated between t and $t + \mathrm{d}t$. Hence one could interpret the time derivative of the relaxation

function as the local probability density of time associated with the formation of the stationary profile at a given location. Then the mean time obtained from this distribution

$$\tau(r) = -\int_0^\infty t\,\frac{\partial R(r,t)}{\partial t}\,\mathrm{d}t \tag{2.4}$$

provides a timescale for the accumulation of the stationary profile at a given location and it's called local accumulation time. Adapting the expression found in [4] to our specific case results in

$$\tau(r) = \frac{1}{2\eta}\int_0^{r\sqrt{\eta/D}} s\left(\left(\frac{\mathrm{K}_1(s)}{\mathrm{K}_0(s)}\right)^2 - 1\right)\mathrm{d}s, \tag{2.5}$$

Since our system has a cutoff radius introduced in order to use periodic boundary conditions, one could think of using as new timestep a value equal or higher to the local accumulation time corresponding to $r = \ell/2$. In this way the stationary condition would be a good approximation. But integrating numerically (2.5) at such value of the radius gives $\tau \sim 0.315$ hours, which seems a big timestep. A compromise could be achieved introducing a new cutoff radius in the simulations and looking to the value of such r_{cut} which is able to produce data the more similar as possible to the ones of Figure 1a. I did exactly that, and the results of this test are presented in Figure 2b, where I used different values of r_{cut} in units of $\sqrt{D/\eta}$. Here it's possible to see how using $r_{cut} = 5\sqrt{D/\eta}$ you seem to obtain a curve almost identical to the initial case. Then computing the local accumulation time at this new distance from the source, you have that $\tau(r)$ is more or less one tenth of the one for $\ell/2$ (Figure 2a). So, taking for example $\mathrm{d}t = 0.04$ hours as timestep, which is what was done from this point onward, one can safely assume the stationary solution to be correct.

2.2.2 The Hill function

There is also another crucial point. In the article the quorum sensing is modeled in a on/off fashion: transtions to the aggregating state are possible only if the local signal's density remains above the sensitivity threshold θ. But it's possible to describe the activation of receptors in a more realistic way. Let's consider an equilibrium reaction between n molecules of M that bind to the

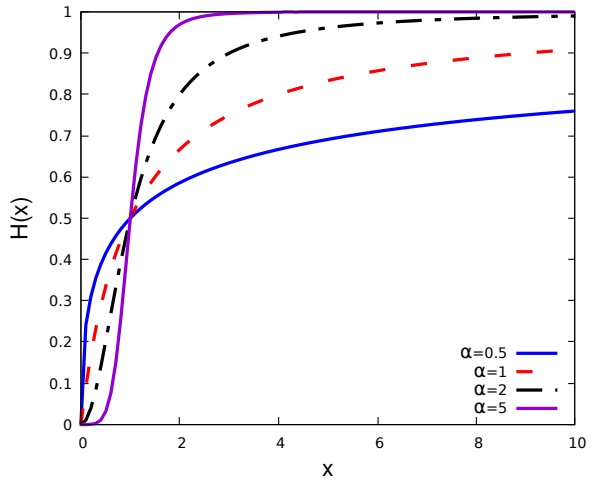


Figure 3: *Plot of the Hill function for different values of the coefficient α and using $\theta = 1$.*

same receptor L:

$$\mathrm{L} + \mathrm{nM} \rightleftharpoons \mathrm{LM}_n$$

Using the molar concentrations of the molecules involved, it's possible to the define the dissociation equilibrium constant relative to the reaction:

$$K_D = \frac{[L][M]^n}{[LM_n]}, \tag{2.6}$$

then the equilibrium ratio of molecules L which have reacted is

$$y_n = \frac{[LM_n]}{[LM_n] + [L]} = \frac{[M]^n}{K_D + [M]^n}. \tag{2.7}$$

Thus from these observations one could substitute the Heaviside function, as a mean to model cells activation, with the Hill function, defined as:

$$H(x) = \frac{(x/\theta)^\alpha}{1 + (x/\theta)^\alpha}, \quad x \in \mathbb{R}^+, \, \theta > 0 \text{ and } \alpha > 0 \tag{2.8}$$

which is indeed a function that goes smoothly from 0 to 1, and approaches the Heaviside func-

tion as $\alpha \to \infty$ (Figure 3).

While the parameter θ is easily interpreted as the value of concentration of signal molecules x at which $H(x) = 1/2$, the Hill coefficient α has been shown to not actually represents the number of binding sites in the experiments. As specified in [5], this coefficient is usually less than the number of binding sites except in case of sequential cooperative binding[2]. So in the end, giving an estimate of α does not allow you to predict the number of sites, but it may tell you something about the cooperativity of the binding sites.

[2]Two ligands B bind cooperatively a receptor A if the binding of the first ligand to an available site increases the chances that a second ligand binds an empty nearby site.

3 Some results from the simulations

3.1 The effect of decreasing cooperativity

While in the original model every cell had a local transition rate $\tilde{\lambda}(\mathbf{r}, t) = \lambda\Theta(\sigma(\mathbf{r}, t) - \theta)$, we are now replacing the Heaviside function with Equation (2.8). The introduction of the Hill function, as one could expect, lowers the loners density as long as the Hill coefficient α is small enough. Indeed the rapidity with which the Hill function goes from 0 to 1 increase with α. This means that now also when the concentration of signal is not so high there's a chance of P-to-A transitions different from zero. Two examples of this effect are presented in Figure 4. Here each point corresponds to the average over the loners densities obtained in the region of the plateau, which in turn are averages over 10 trials, as done previously. The curves, for the two values of K considered, start for $\alpha = 1$ at a very low value with respect to the Heaviside case (corresponding to the horizontal red line) and then approaches the latter as the Hill coefficient grows.

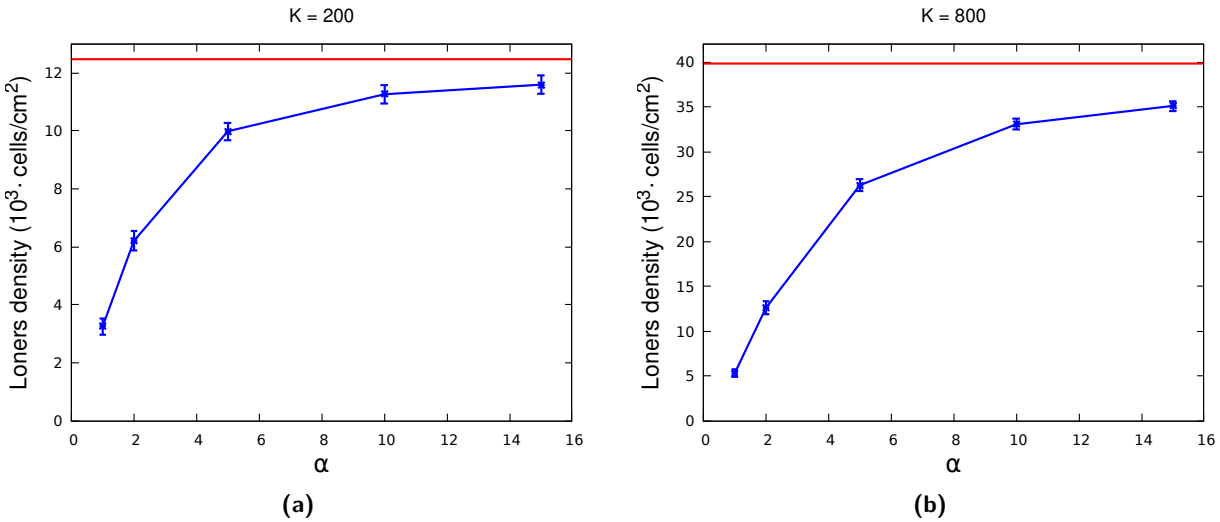
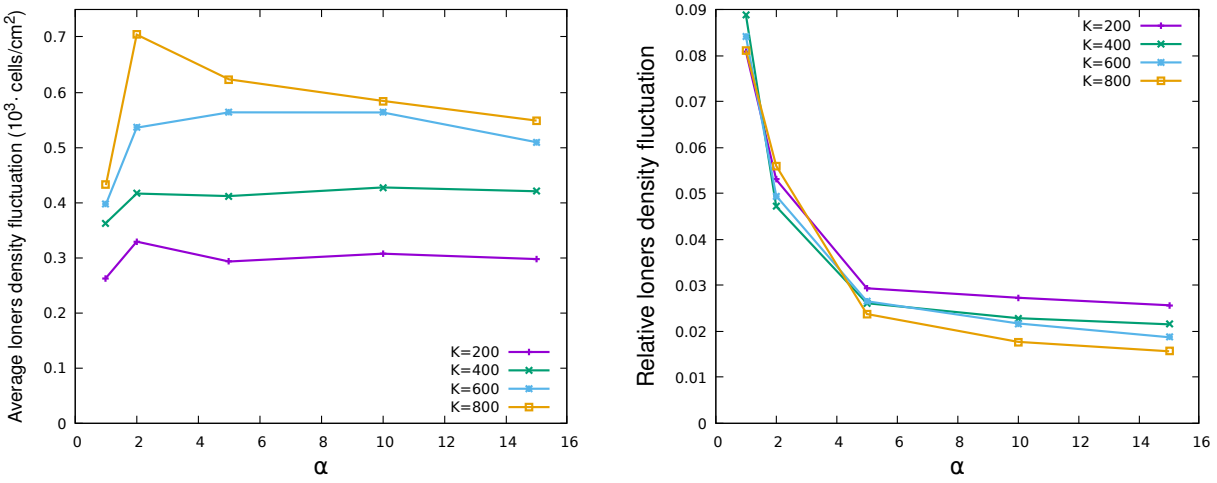


Figure 4: *Average loners density in the region of the plateau, obtained using different values of the Hill coefficient. The error bars correspond to the average standard deviation, while the red line is the value obtained using the Heaviside function.*

Similarly, the plots in Figure 5a show how the fluctuations between trials are affected. As for the precedent figure, also here every point is the average of the standard deviations computed over 10 trials, obtained from various initial density of plated cells in the region of the plateau. It is possible to observe how, apart for the case $\alpha = 1$, the value of the fluctuations decrease with α until reaching saturation, indicating that the Hill function affects the distribution of loners not only in the mean value but also with respect to the variance.



(a) *Average standard deviation of loners' density as a function of the Hill's coefficient, for different values of $K = \theta/\gamma$.*

(b) *Average standard deviation divided by the mean loners' density, as a function of the Hill's coefficient and for different values of $K = \theta/\gamma$.*

Figure 5: *Every point represents the average between standard deviations corresponding to different values of the initial cell density, in the region of the plateau.*

Thus the fluctuations seem to behave similarly to the value of the plateau, in the sense of tending asymptotically to the Heaviside value, but with the difference that deacresing the Hill coefficient you obtain an higher average standard deviation (apart from the case $\alpha = 1$). In particular the relative fluctuations do not show a precise dependence on K for very low values of α, while as the Hill function becomes more similar to $\Theta(x)$ we have that it decreases with θ/γ, which is what happens in the Heaviside case.

These effects are connected with the fact that in principle, as already said, waiting enough time it's always possible to have new transitions since the rate $\tilde{\lambda}(\mathbf{r}, t)$ is not exactly zero when the signal density is very low. One expect this to be especially true for $\alpha = 1$, value at which the Hill function loses its sigmoid-like shape. But if this is the case, then the natural consequence is the inevitable dependence on t_{arr} of the value of the plateau, dependence which of course should be stronger in the $\alpha = 1$ case. Figure 6 confirms to us exactly that: a higher t_{arr} leads to a lower final density of loners, whose relative difference with the case $t_{arr} = 1$ hour increases for low α and high t_{arr}. This phenomenon it is not observed using instead the Heaviside function.

But this raises inevitably a crucial question: has the concept of loners any meaning at this point? Clearly while in theory under these conditions the loners density should go to 0 as the time $t \to \infty$, we're dealing with biological systems, which have limited lifespan and resources. At a certain point the transition probability should become so tiny that is not able to change the loners' density in a realistic timescale for living organisms. Indeed Tarnita *et al.* noted that in their experiments the aggregation never lasted longer than 22 hours: at a certain point the mound enters the slug phase and it starts to travel away from the aggregation site, eventually absorbing some loners. Since using $t_{arr} = 1$ hour results already in simulations longer than the aforementioned limit, especially for low values of α, I decided to keep such value.

3.2 Spatial heterogeneity

To understand better the above point one could ask how the signal concentration or the Hill function evolves during the simulation. In order to explore that, I implemented the possibility for the code to store the average signal density corresponding to a given simulation time. Such average concentration is computed using the local signal density seen by each P-cell, hence it is an average whose population is changing with time. Additionally it also stores the corresponding average Hill

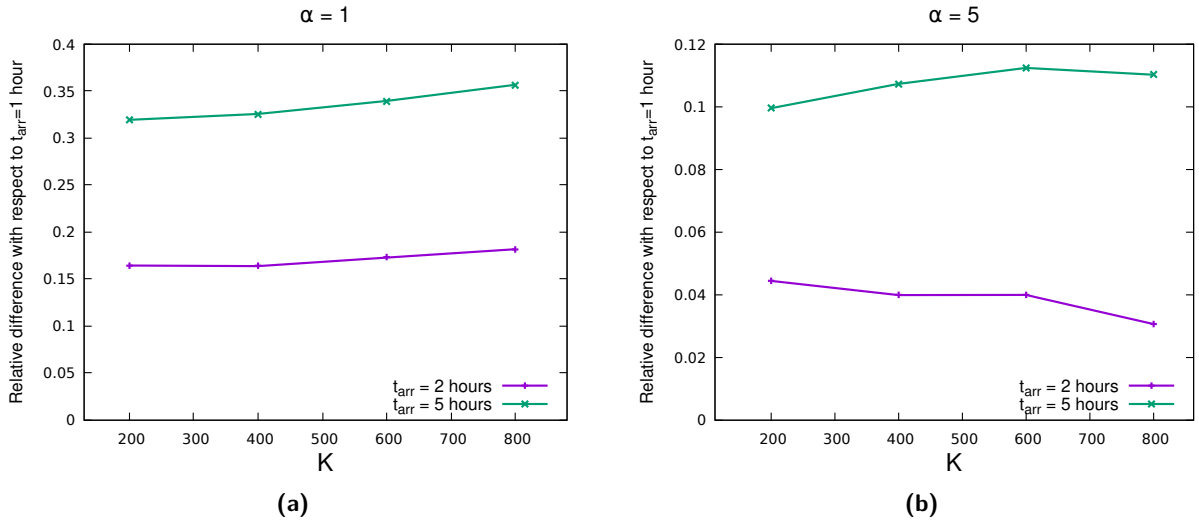
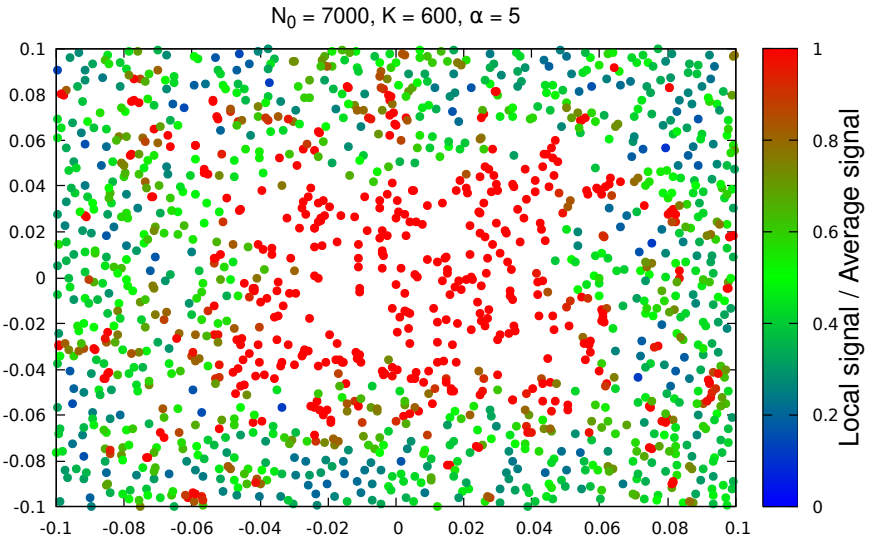


Figure 6: *Plots of the relative difference of average loners' densities between the* $t_{arr} = 1$ *case and the case with a different* t_{arr}.

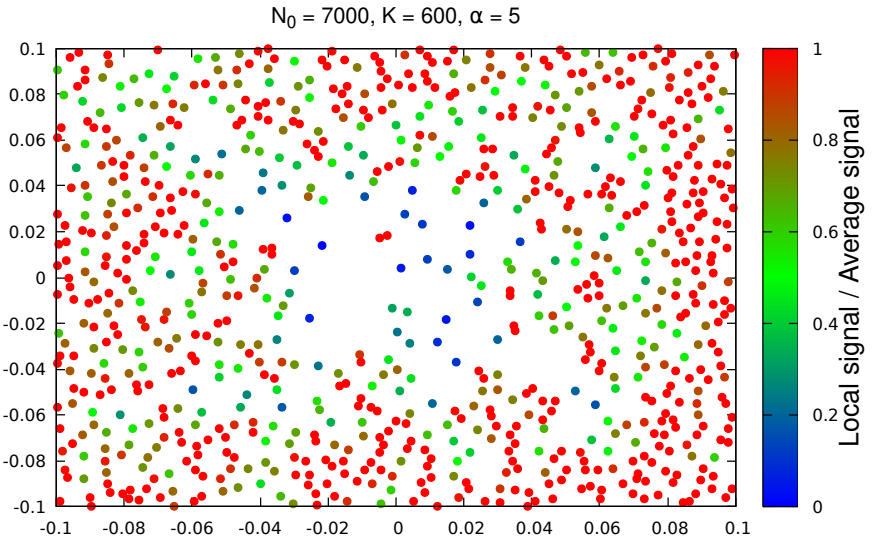
function, number of loners and for certain times it takes a snapshot of the simulation, storing the position and local signal density at that very same location for every P-cell present in that given instant.

Such instants are the initial and final time, plus the moment after which $N_0/2$ cells converged into the mound (Figure 7). What can be seen is how the signal is on average higher around the aggregation center during intermediate times, while the situation is reversed at the end of the simulation, where is the external region which contains an higher density. This happens thanks to the A-cells, which travel towards the center of the tile causing the signal concentration to grow in this zone. This makes easier the occurrence of P-to-A transitions, which in turn leads to a lower final density of amoebae, and thus signal, around the center.

It is possible to observe distributions with the same features as in Figure 7 also using the Heaviside function. The difference with the Hill case is that of course it is not present an influence from α. In fact changing K and α not only affects the final density of loners (and signalling molecules)



(a) P-*cells population for the instant at which half of the initial number of cells have become* M-*cells.*

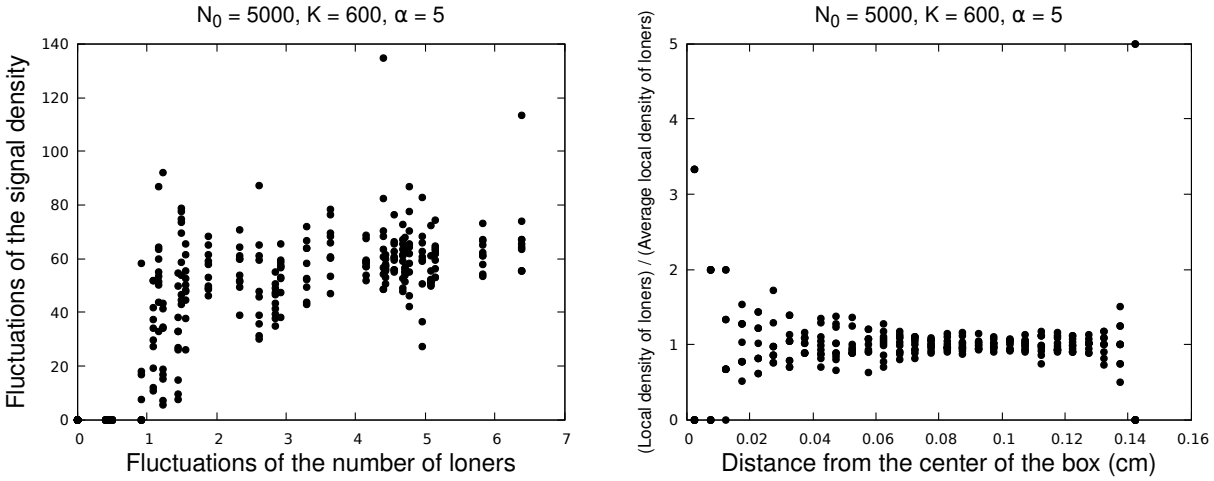


(b) P-*cells population for the final instant.*

Figure 7: *Relative signal density and position for the population of amoebae in the P-state at intermediate* ***(a)*** *and final time* ***(b)****. Some points have actually a ratio between local and global concentration of signal higher than 1, but the scale was set in this way in order to have more clarity.*

but also the dimension of the central low density region.

We can use these snapshots to investigate how loners and signal densities behave as functions of the distance from the center of the tile. I divided my square box in concentric rings and for each one I counted the number of loners inside and computed the average signal density and its standard deviation. An interesting feature that came out is the relationship between the local fluctuations of signal and the local fluctuation of loners, because it actually seems that there is almost no relationship at all.



(a) *Local fluctuations of the signal density as a function of the local fluctuations of the number of loners.*

(b) *Ratio between the local loners density and its average over trials, as a function of the distance from the center.*

Figure 8: *These plots contain data from 10 different trials. All the quantities shown were computed as functions of the distance from the center of the square box.*

The data produced are showed in Figure 8a. Here the fluctuations of the number of loners were computed, as usual, over 10 different trials, while the local standard deviations of the signal density for all these trials are reported. As it can be seen, apart from some points, while the fluctuations of signal density remain around an average value, the fluctuations of loners keep growing. In particular the points where there's not this trend corresponds to the locations with fewer loners,

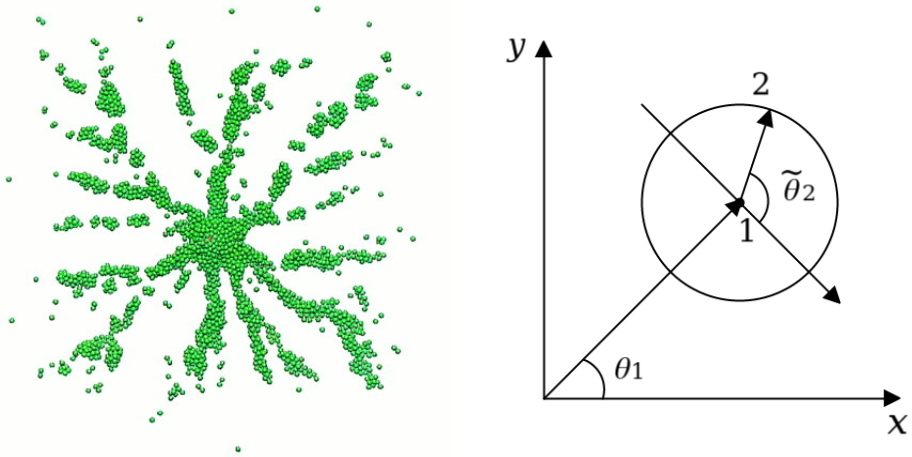
or in other words distances very close to the center of the box, hence in the "emptied zone", and $\ell/2 < r < \sqrt{2}\ell/2$, where the considered rings intersect only partially with the square box.

Another thing that can be noted is that while the external region seems to contribute more on the fluctuations of loners, since it contains the vast majority of the remaining amoebae, this is also the zone with the less variability between trials as opposed to the region near $r = 0$ (Figure 8b). Indeed, for distances around $\ell/2 = 0.1$ cm, it can be seen that the local number of loners assumes values very close to each other for different realizations. The same feature can be found also looking at the local density of signal.

3.3 A possible rule for anisotropic interactions

A way that could hopefully lead to the reproduction of the strain variability observed experimentally would be to try to break this homogeneity between trials in the outer region. For example introducing some anisotropy in the system. This is not completely far fetched. In fact it results from experiments that in real systems, during the aggregation, the amoebae travel initially by themselves, but at later stages they form connected streams that migrate toward the aggregation center [6]. In order to mimic the formation of these streams I introduced a simple interaction between A-cell that had the effect of making them align while travelling toward the center.

More spefically, starting from the system of coordinates having as origin the center of the square box, if we consider an amoeba, which we'll call amoeba 1, that forms an angle θ_1 with the x-axis and has as nearest neighbour the amoeba 2, then the presence of 2 introduces a modification in the velocity of 1, and viceversa. In order to obtain this new value I first need to go in the system of reference of the amoeba 1 and make a clockwise rotation of an angle $\varphi = \pi/2 - \theta_1$. In this way my new y-axis is parallel to the vector connecting the center of the tile to the position of 1. Calling $\tilde{\theta}_2$ the angle that the amoeba 2 forms with the new x-axis (Figure 9b) then you can define the velocity



(a) *Simulation of the aggregation pattern obtained in [6] using a more realistic model for the cell movement in D. discoideum.*

(b) *Graphical schematization of the quantities involved in Equation (3.1).*

Figure 9: *Plots relative to the discussion about the anisotropic interaction.*

with which the amoeba 1 travels as

$$\mathbf{v} = v_{\parallel}\mathbf{u}_r + v_{\perp}\mathbf{u}_{\perp}, \quad \begin{cases} v_{\parallel} = -v\sqrt{1-\xi^2\cos^2\tilde{\theta}_2} \\ v_{\perp} = v\xi\cos\tilde{\theta}_2 \end{cases} \tag{3.1}$$

where ξ is a parameter between 0 and 1, which determines how much of the module v (that is kept constant) is going to the perpendicular component $v_{\perp}$. In this way the cells align themselves while traveling toward the center of the box: for example, if cell 2 is "at the right" of cell 1 (with respect to the radial direction) then $\cos\tilde{\theta}_2$ will be positive, while from the point of view of 2 we will have that $\cos\tilde{\theta}_1$ is negative, thus creating an attraction. This interaction makes the cells slow down in the radial direction mainly when they are not aligned, which for sure does not incourage the overlap of amoebae but can in some occasion cause it. In order to deal with this I imposed that when the distance between the two cells goes below a certain (small) value the interaction is turned off. This makes the two cells move like a unique block that can only attract surrounding cells and it is not attracted by others.

Unfortunately the effects on the fluctuations of signal and loners where minimal and did not produced what hoped for.

4 The time evolution of the loners

As introduced before, in order to understand better the effects on the dynamics produced by the presence of the Hill's function, I made different simulations in which I stored the time evolution of various quantities of interest. As a tool of analysis I introduced an empirical expression capable of describing such evolution, and proceeded to fit my data through different parameters.

In particular at the beginning I was interested on the relationship between the average global signal $\bar{\sigma}(t)$ and the average Hill function $\bar{H}(t,\alpha)$ (where now I'm making explicit the dependence on α). After observing that the two quantities are not related trivially,[3] I realized that the data describing the time evolution of the average Hill function and the number of loners produce curves qualitatively very similar. So I decided to focus my efforts on describing the dynamics of the population of P- and A-cells, which at the end of the simulation will correspond to the P-cells that did not transitioned.

Influenced by the previous steps, I defined also in this case an empirical function, which is essentially a modified Hill function. Clearly one could argue that it is not obvious that the number of loners have such a time dependence. We are basically dealing with sigmoid curves, any other function with the same qualities would have been in principle equally valid. This is indeed true, but this expression has proven to be quite effective, for this reason I decided to keep it.

The function that I used is:

$$L(t) = \begin{cases} (N_0 - D(K,\alpha)) \cdot (1 - h(t)) + D(K,\alpha) & \text{if } D(K,\alpha) \leq N_0,\ t^*(K,\alpha) > 0 \\ N_0 & \text{otherwise} \end{cases} \tag{4.1}$$

[3] More precisely I checked if $\bar{H}^{em.}(t,\alpha) = H(\bar{\sigma}^{em.}(t),\alpha)$, where the "$em.$" apex indicates an empirical expression. But as it turns out you don't have such a simple relationship between the two average quantities.

where

$$h(t) = \frac{\left(t/t^*(K,\alpha)\right)^{A(K,\alpha)}}{1+\left(t/t^*(K,\alpha)\right)^{A(K,\alpha)}} \tag{4.2}$$

Actually I should have written $A_{N_0}(K,\alpha)$ etc., because it is also present a dependence on the initial number of cells, but in the following I'll omit that in order to keep the notation light. Other than that, I'm completely ignoring the effect of the ratio λ/v since I will keep it fixed.

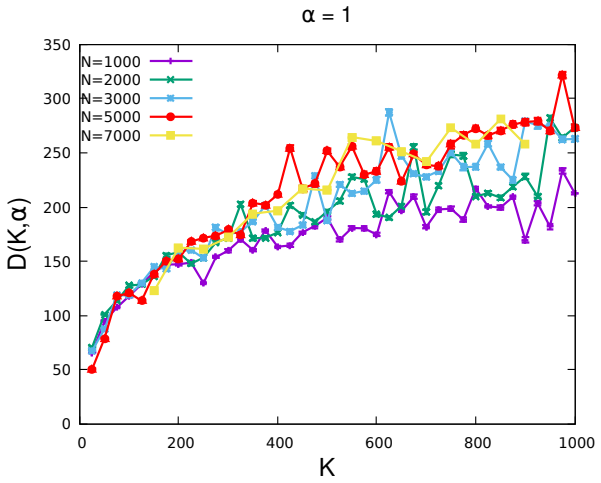
4.1 The parameters of the fit

I will present only what happens to the parameters of the fit varying K, because the dual case (keeping fixed K and varying α) it's not interesting: you just have a saturation of the various parameters towards asymptotic values in a way similar to what seen for the plateau (Figure 4). This is clearly unsurprising since it just represents the convergence to the Heaviside function.

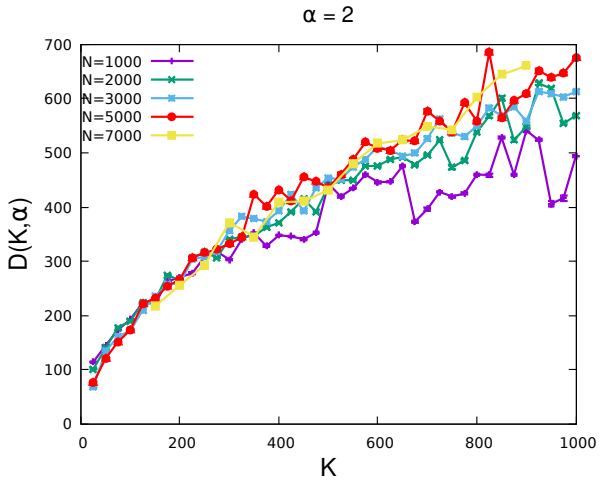
We can start our analysis with the parameter which has the most intuitive interpretation: the baseline $D(K,\alpha)$, which of course corresponds to the final number of loners. Indeed in the limit of $t \to \infty$, you have that $h(t) \to 1$ and thus $L(t) \to D(K,\alpha)$.

Taking a very high value of the Hill coefficient like in Figure 10d, where the data corresponding to $\alpha = 60$ are presented, we obtain a situation essentially equivalent to having an on/off quorum sensing. This can be seen using the same code to store the time evolution in the case of the Heaviside function. Clearly even if the Hill coefficient is very high this correspondence is not perfect, but the agreement with the data is very good (data not presented in this text).

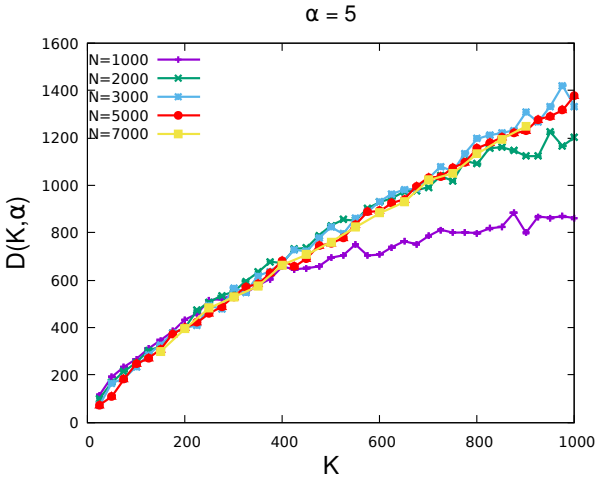
Always remaining on Figure 10d, we can see that the final number of loners grows initially in a linear fashion, representing the reaching of the plateau: as we have seen, as long as N_0 is high enough and the ratio $K = \theta/\gamma$ not too big, the value at which the population of remaining P-cells



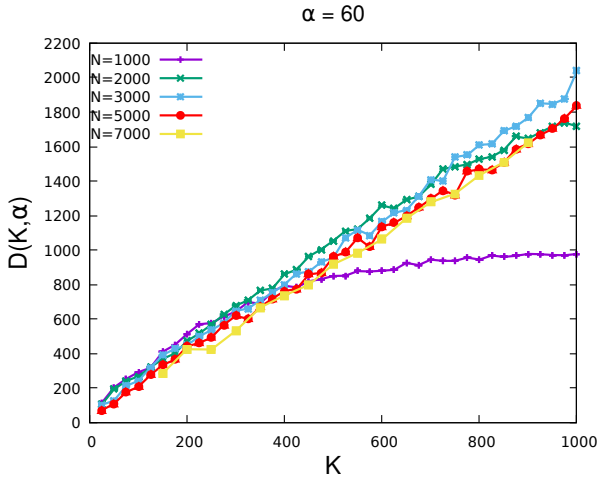
(a) *Baseline of $L(t)$ for $\alpha = 1$.*



(b) *Baseline of $L(t)$ for $\alpha = 2$.*



(c) *Baseline of $L(t)$ for $\alpha = 5$.*



(d) *Baseline of $L(t)$ for $\alpha = 60$.*

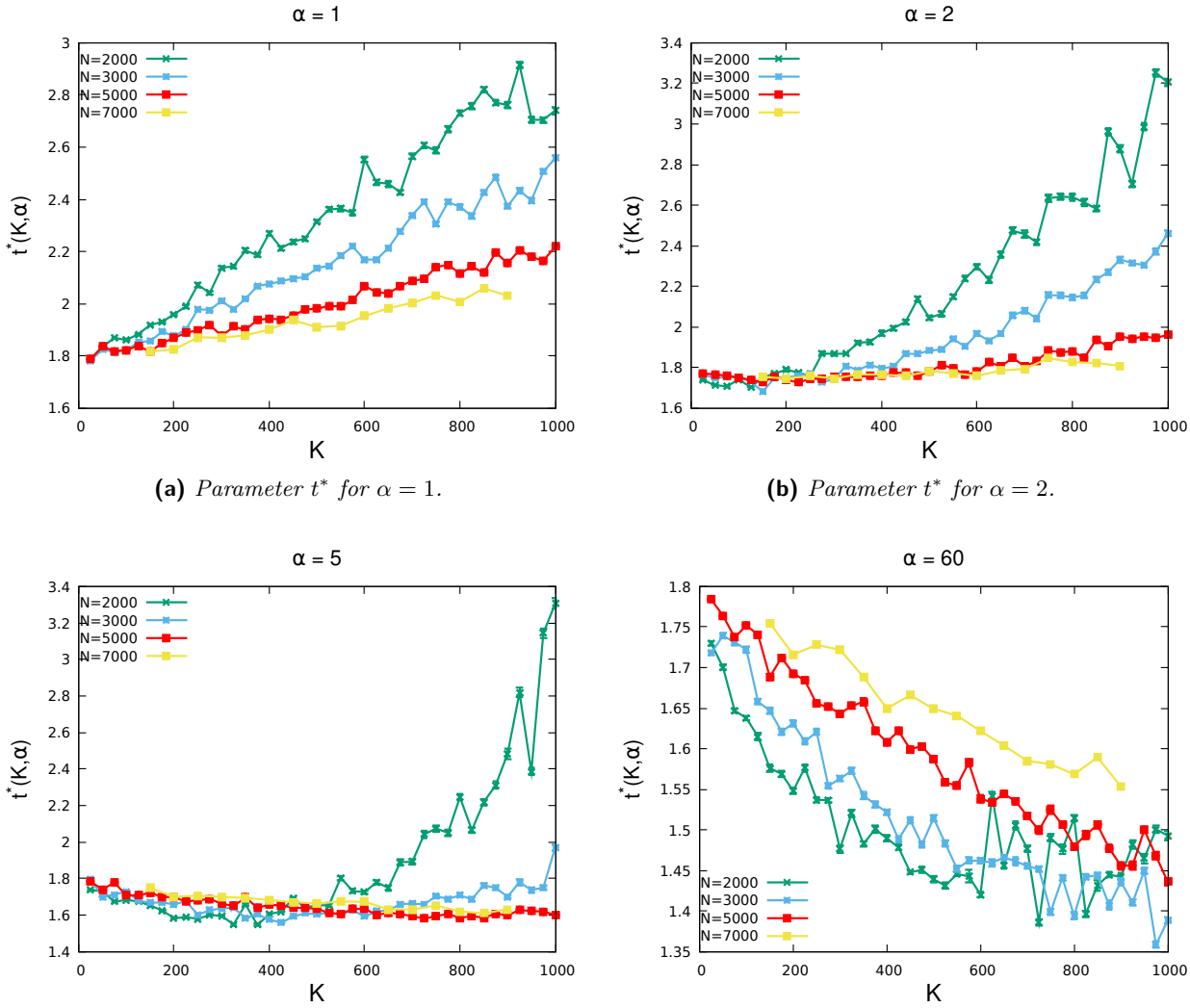
Figure 10: *Plots containing the data relative to the D parameter obtained from the fits for different α and N_0, and as a function of K. The error bars are given by the error of the fit. The curves corresponding to $N_0 = 7000$ contain less points since their simulations are more time consuming.*

saturate does not change with N_0 and grows with K (Figure 1a). Indeed by looking at these points it can be noted that they seem to be weakly dependent on N_0. Then $D(K, \alpha = 60)$ starts to bend for higher K's and saturates around N_0 (as it's clear for the curve corresponding to $N_0 = 1000$) indicating that there's little or not activity at all.

This kind of behavior is true also for lower α's, only here, other than having noisier curves, the "linear phase" is more bent, resembling something $\sim x^p$, with $0 < p < 1$. It seems that in some way there's an anticipation and a slowing down of the beginning of the saturation toward the baseline. In fact while for $\alpha = 60$ the $N_0 = 1000$ curve bends around $D \sim 800$, for $\alpha = 5$ the same curve detach from the group around $D \sim 600$ (Figure 10c). Another example of this behavior can be seen looking at the $N_0 = 2000$ curve, which seems to be initiating saturation at very last two points of its curve in Figure 10d, while in the other plots it begins earlier. But this does not mean that if for given N_0 and K I'm inside of the plateau zone, lowering α one can in general go outside of it, but just the fact that the saturation towards N_0 is slower. Moreover, making simulations for higher K's (not presented here) it can be verified that indeed these curves continue their slow growth until reaching the value N_0. This is in agreement with what we've seen before: the Hill function facilitates aggregation, lowering the value of the plateau. Thus having a simulation in which no amoeba arrives to the mound becomes more difficult as α is decreased.

Let's now focus instead on $t^*(K, \alpha)$: this parameter corresponds to the instant at which $L(t^*) = (N_0 + D(K, \alpha))/2$, hence the time needed for half the final M-cells population to arrive to the aggregation point. Some results from the various fit are presented in Figure 11, where I excluded the data relative to $N_0 = 1000$ since such curve explodes very quickly. In fact as can be easily observed, many t^* curves seems to grow toward infinity as $K \to \infty$, and such growth seems stronger for lower numbers of initially plated cells.

Using higher K's what can be observed is that you obtain a longer time t^*, until you have a



(a) *Parameter t^* for $\alpha = 1$.*

(b) *Parameter t^* for $\alpha = 2$.*

(c) *Parameter t^* for $\alpha = 5$.*

(d) *Parameter t^* for $\alpha = 60$.*

Figure 11: *Plots containing the data relative to the t^* parameter obtained from the fits for different α and N_0, and as a function of K. The error bars are given by the error of the fit. The curves corresponding to $N_0 = 7000$ contain less points since their simulations are more time consuming..*

discontinuity in its value. Usually it becomes suddenly very small, but it can also happen the opposite for the case in which t^* saturates (which will be discussed below). In practice what happens is that the simulation has reached a point in which the time evolution of the loners is represented by a stepwise curve composed by a relatively small number of jumps. The value of the threshold is so big (or alternatively the production rate γ is so small) that very few amoebae arrive to the aggregation point, until, increasing K, you do not register any activity at all. Clearly this does not happen immediately, but instead in a relatively short window of K-values. Additionally you can obtain the same effect increasing the Hill coefficient: if the ratio between the density of signal and θ is small (which is the case for high K), by increasing α you are reducing the probabilty of P-to-A transition per unit time. Since these scenarios are not the focus of the model, I will neglect such data in my analysis and just consider them as $L(t)$ remaining at the initial value N_0 at any time.

Returning to Figure 11, it can be noted that while for $\alpha > 1$ you have an initial decrease (which in the plots showed is extremely clear for $\alpha = 60$, but present also in the other two), indicating a sort of parabolic dependence on K, for $\alpha = 1$ this does not happen, representing an exception.

Adopting once again our approximation of identifying the data for $\alpha = 60$ with the Heaviside case, this change in the first derivative's sign could be interpreted as the point where you cross the region in which changing N_0 you obtain the same final number of loners, or in other words there's a plateau, and the region where this does not happen. Indeed confronting Figure 11d with Figure 1a this seems the case. Keeping in mind that $N_0 = 2000$ cells corresponds to $5 \cdot 10^4$ cells/cm^2 and $N_0 = 3000$ to $7.5 \cdot 10^4$ cells/cm^2, this is coherent: the latter's curve stops decreasing and stabilize around $K \sim 800$, while the former at $K \sim 600$. Thus one could be tempted of thinking that $t^* \to \infty$ represents the situation in which it takes an infinite time for the population to reach the value $(N_0 + D)/2$, and hence representing the case described previously of $L(t) = N_0$. The problem with this interpretation is that going back to the plots relative to lower value of the Hill coefficient you observe some contradictions. In fact in this way for $N_0 = 5000$ you have that $K = 800$ is in

the plateau for $\alpha = 5$ and higher, but not for $\alpha = 2$. What's worse is that for $\alpha = 1$ you would find yourself always outside of the plateau zone. In addition, in Figure 11d the curves for the lower value of N_0 seem to stop growing after a certain point (which can be confirmed using higher values of K).

Finally for what concern $A(K, \alpha)$, its behavior has the same features of $t^*(K, \alpha)$, only with opposite convexity.

4.2 The interpretation of t^*

In order to shed some light on the strange behavior of t^* I repeated the same fitting procedure for the time evolution of amoebae which have a transition rate governed by the Heaviside function. What I've found is qualitatively in line with Figure 11d, with the crucial common detail that $t^*(K)$, once that it has stopped decreasing, starts to oscillate around a certain N_0-dependent value, for K high enough. This is in contraddiction with what happens for the rest of the plots in Figure 11: indeed if the Hill coefficient is small enough, increasing K, it's possible to observe an increase of t^* for the various curves.

But then if this is the case what causes the growth? It's possible to answer to this question by looking at the time evolution itself of the loners population (Figure 12). At the beginning of the simulation, in a situation in which α is finite, N_0 is high enough and K relatively low, the signal's density felt by the cells is high and the Hill function is of the order 10^{-1}. The first A-cells start traveling towards the center of the box, making the global density of signal lower in the external regions and increasing it around the aggregation point. As a consequence you have further transitions in this zone, causing a mass arrival to the mound. For this reason you'll have a lag in the response of the population to the occurence of the transitions (since we are counting the amoebae which are not M-cells), followed by a fast change.

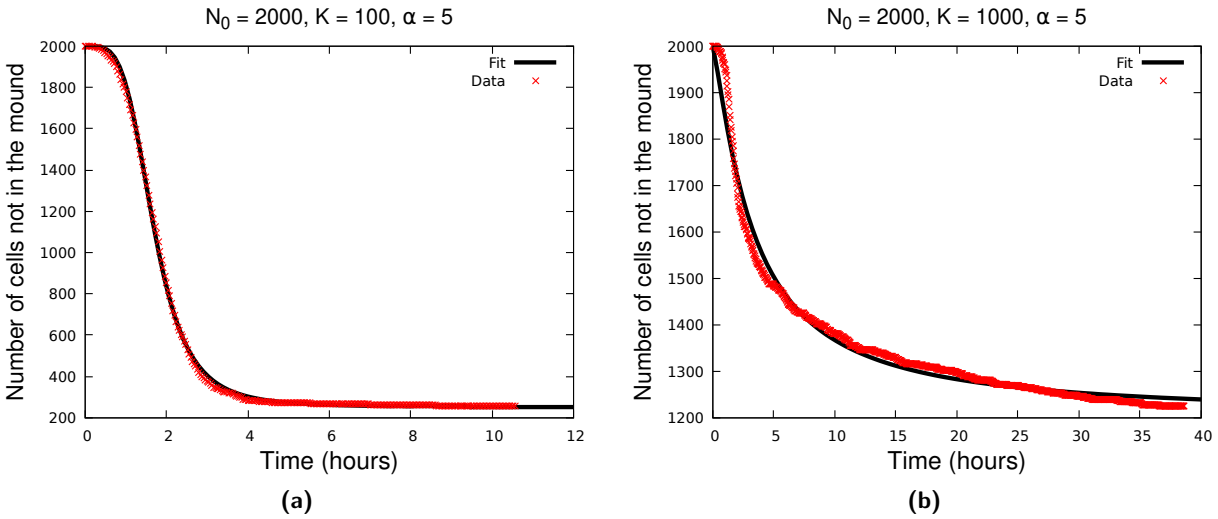


Figure 12: *Time evolution of the population of amoebae outside the mound, together with the empirical curve* $L(t)$.

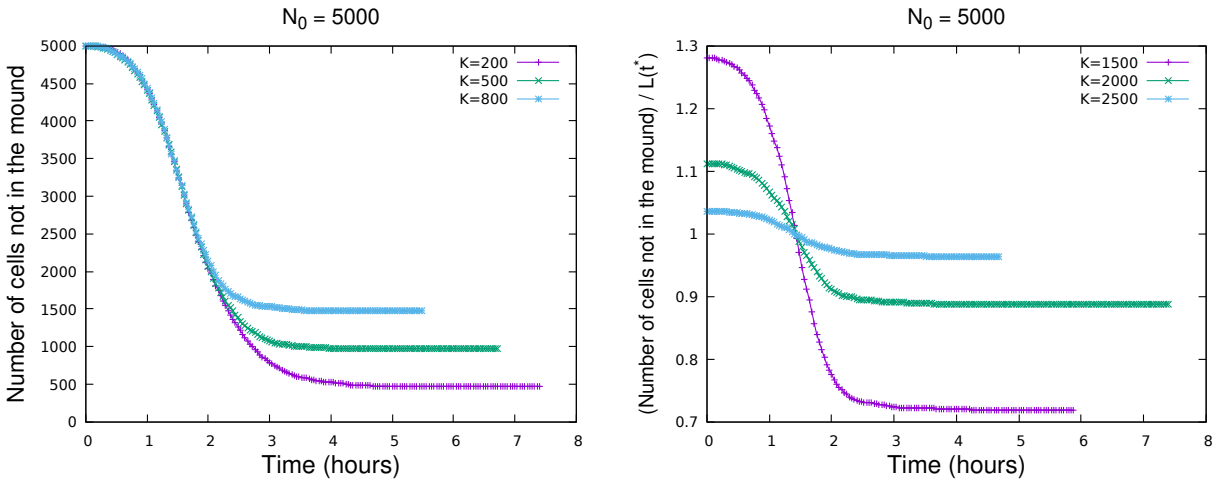
What described until now is what happens in Figure 12a, where it can be seen how the first points remain around the initial value $N_0 = 2000$. Then you have a relatively fast decrease, containing the vast majority of the arrivals to the mound. From there the population of loners goes down in a relatively slow fashion, representing a situation where the signal's density is on average very low, except in some location where there are still few transitions and thus arrivals. So we can schematize the dynamic dividing it in three phases: a "lag" phase, a "burst" phase and a "slow-decreasing" phase.

These features appear going toward low K and high α and are in common with the Heaviside case. Indeed a large Hill coefficient means a high first derivative of the Hill function around the point in which the signal is equal to the threshold θ. Hence a small variation of the density of signal can take you from a situation in which the transition rate is high, to one in which is almost zero. At the same time, if the ratio θ/γ is small, you have more easily $H(\sigma(\mathbf{r}, t))$ closer to one.

The specular direction of high K and low α instead have the effect of making the third slow-

decreasing phase more relevant in the dynamics. Figure 12b is one example of such behavior, the global time for which the simulation last is much longer and in general the curve loses its sigmoid shape. This is in principle a serious problem, becuase the function $L(t)$ was thought to describe a type of dynamics like the one in Figure 12a, making the growth of t^* to appear as a spurious effect.

Indeed if one wants to describe better the time evolution of loners it is needed the introduction, for example, of another component in $L(t)$ which decreases slowly with respect to the Hill function and which will be negligible for high α. While it is true that the empirical curve in Figure 12b does not describe the time evolution as well as in the case of $K = 100$, nonetheless it contains informations about the influence of the Hill function and does not represent a bad approximation. Besides the growth of t^* assumes now a precise interpretation: it is the sign of a time evolution of the loners population deeply affected by the presence of a non-zero transition rate also for small concentration of signalling molecules. For these reasons I've decided to keep this expression, even if imperfect.



(a) *Curves relative to values of K inside the plateau zone.*

(b) *Curves relative to values of K that do not correspond to a plateau situation for $N_0 = 5000$.*

Figure 13: *Time evolution of the population of amoebae outside the mound for the Heaviside function.*

Using what it's been said until now, we can finally shed some light on the observed behaviors.

For example, the similarity with the Heaviside function of the low K case can help us understand why in Figure 11 you have that t^* goes down for $\alpha \neq 1$. In fact we know that in the original model, when the formation of the plateau is possible, the amoebae continue to aggregate until the signal concentration reaches everywhere a value under the threshold θ. A certain number of cells is left behind, and in particular this number grows with the ratio θ/γ. This means that the time at which half of the final M-cells population has arrived to the mound has to depend on the final number of loners. Or in other words, if D goes up when you increase K, you need less time to arrive at halfway the baseline. This can be understood by comparing the time evolution of different values of K (inside the plateau zone) for the Heaviside case (Figure 13). Indeed the different curves represented in Figure 13a have the same time evolution for the points away from their respective final number of loners. Such feature is lost going outside of the plateau, where is instead possible to observe the aforementioned stationarity of t^* for large K's. This can be seen dividing the number of amoebae outside the mound by $L(t^*) = (N_0 + D)/2$, as done in Figure 13b. What we obtain is that the different curves reach 1 more or less at the same time.

But why does this happen? When K is high enough to make the plateau disappear, but not so big that $L(t) = N_0$, the signal density felt by the cells is enough to trigger only a relatively small number of transitions. Further events can only happen (for the vast majority) in the time window represented by the lag phase, during which, as said before, the A-cells travel toward the center of the square box creating additional opportunities for the amoebae around the aggregation point to transition. The suspect is that t^*, that can now also be understood as the lag time, is proportional to the characteristic transition time $1/\lambda$, since it depends heavily on how much amoebae transitioned at the beginning of the simulation. But clearly it is also present a dependence on v, since of course faster cells need less time to reach the center.

The stationarity of t^*, as seen, is present also in the Hill case and starts to appear for very high α's. For smaller values instead we observe t^* changing slope, meaning that the dynamics is

influenced by the modifications to the model. But this makes us understand that for $\alpha = 1$ the third slow-decreasing phase is always important: you have strong effects due to the Hill function for every K.

It follows that knowing the expression of $t^*(K, \alpha)$ can give you informations on whetever you are in the plateau or not, and on the influence of the Hill function on the process. This is important because it's something we can measure in real-life experiments, together with the final number of loners, described by $D(K, \alpha)$. Thus by monitoring the time evolution of the aggregation it's in principle possible to obtain an estimate of the ratio $K = \theta/\gamma$ and the Hill coefficient α, given that one has an expression for such parameters. Since the molecule responsible for the quorum sensing process of *Dictyostelium discoideum* is still unkown, knowing α could be a precious hint in identifying such signalling molecule.

4.3 Conclusions and Perspectives

An empirical expression for the parameters in Equation (B) that reproduces well the data from the simulations, alongside the features underlined in this last section, can be found. In Appendix B I reported an expression that has proven to work quite well for $K \in [250, 1000]$, but such function reproduce the data from the simulations only for $N_0 = 5000$. In fact in order to achieve its purpose of helping in the identification of the signalling molecule it would be necessary an amount of data bigger than the one produced. As said before there are also present dependences on the ratio λ/v and N_0, which were not explored at all, requiring additional simulations to be made. Indeed while it could be in theory manageable during an experiment with real systems to control the initial number or density of amoebae, it is of course not the same for the other variable. Such dependence have to be taken into account and how it is affected by the Hill function is yet to be studied. In addition the big mass of data that one would obtain exploring the effects of changing λ/v and N_0 (alongside K and α) would require the implementation of some more sophisticated tecniques for

the analysis, like a machine learning algorithm.

A Appendix: the non-stationary solution

In order to obtain the non-stationary solution, I rewrite the generalized diffusion equation describing the concentration of signal emitted by a single cell in the following way:

$$\frac{\partial\sigma(x,y,t)}{\partial t} = D\nabla^2\sigma(x,y,t) - \eta\sigma(x,y,t) + S(x,y,t)$$

where I've introduced an additional term representing the source. In our specific case we are considering punctual sources which start to emit at $t = 0$ with emission rate γ, thus:

$$S(x,y,t) = \gamma\,\delta(x)\delta(y)\Theta(t)$$

Making a Fourier transform both in space and time I obtain the following equation

$$\mathrm{i}\omega\,\tilde{\sigma}(\vec{k},\omega) = -D(k_x^2 + k_y^2 + \frac{\eta}{D})\,\tilde{\sigma}(\vec{k},\omega) + \frac{\mathrm{i}\gamma}{\omega},$$

with solution

$$\tilde{\sigma}(\vec{k},\omega) = -\frac{\gamma}{D|\vec{k}|^2+\eta}\left(\frac{1}{\mathrm{i}\omega} - \frac{1}{\mathrm{i}\omega + D|\vec{k}|^2+\eta}\right)$$

At this point I compute first for simplicity the antitransform with respect to time, which is just the sum of two notable antitransformations:

$$\tilde{\sigma}(\vec{k},t) = \frac{\gamma}{D|\vec{k}|^2+\eta}\left[1 - e^{-(D|\vec{k}|^2+\eta)\,t}\right]\Theta(t)$$

In order to solve also with respect to the spatial coordinates I need first, defining $\beta = \sqrt{\eta/D}$, to compute the antitransform of the stationary component of the solution, represented by the first term. This is actually (neglecting the Heaviside function) a notable Henkel transformation [3]:

$$\mathcal{F}^{-1}\left(\frac{1}{|\vec{k}|^2+\beta^2}\right) = \frac{1}{2\pi}\int_0^\infty \frac{k\,\mathrm{J}_0(kr)}{k^2+\beta^2}\,\mathrm{d}k = \frac{1}{2\pi}\,\mathrm{K}_0(\beta r)$$

Which is exactly, apart for some coefficients, the stationary solution found in [1].

The complete non-stationary solution can be found noticing that the antitrasformation of the non-stationary term can be expressed as a convolution between the expression above and the usual punctual solution of the diffusion equation (with a degradative factor of course):

$$\begin{aligned}\mathcal{F}^{-1}\left(\frac{e^{-(D|\vec{k}|^2+\eta)\,t}}{|\vec{k}|^2+\beta^2}\right) &= \frac{e^{-\eta t}}{4\pi Dt}\int \mathrm{d}x_0\mathrm{d}y_0\,\frac{\mathrm{K}_0}{2\pi}(\beta\sqrt{x_0^2+y_0^2})\,e^{-(\vec{r}-\vec{r}_0)^2/4Dt} = \\ &= \frac{e^{-\eta t}e^{-r^2/4Dt}}{4\pi Dt}\int r_0\mathrm{d}r_0\mathrm{d}\theta_0\,\frac{\mathrm{K}_0(\beta r_0)}{2\pi}\,e^{-r_0^2/4Dt}\,e^{-\frac{rr_0}{2Dt}\cos(\theta_0-\theta)}\end{aligned}$$

It is possible to recognize in the integral above the definition of the zero order modified Bessel function of the first kind:

$$\mathrm{I}_0(s) = \frac{1}{\pi}\int_0^\pi \mathrm{d}\theta\,\exp(s\cos\theta)$$

Using this expression, the convolution can be finally rewritten as

$$\frac{e^{-\eta t}e^{-r^2/4Dt}}{4\pi Dt}\int_0^\infty r_0\mathrm{d}r_0\,\mathrm{K}_0(\beta r_0)\,e^{-r_0^2/4Dt}\,\mathrm{I}_0\left(\frac{rr_0}{2Dt}\right)$$

Finally, putting all together we obtain exactly Equation (2.1).

B Appendix: the empirical expression

In this section I report the formulae obtained for the empirical function describing the time evolution of amoebae outside the mound. More precisely these expressions work for $N_0 = 5000$ and $K \in [250, 1000]$. As seen in Section 4 the function that I'm using is

$$L(t) = \begin{cases} (N_0 - D(K,\alpha)) \cdot (1 - h(t)) + D(K,\alpha) & \text{if } D(K,\alpha) \leq N_0,\ t^*(K,\alpha) > 0 \\ N_0 & \text{otherwise} \end{cases}$$

with

$$h(t) = \frac{\left(t/t^*(K,\alpha)\right)^{A(K,\alpha)}}{1 + \left(t/t^*(K,\alpha)\right)^{A(K,\alpha)}}$$

In particular the function describing the Hill's coefficient is

$$A(K,\alpha) = \begin{cases} mA(\alpha) \cdot K - pA(\alpha) \cdot \sqrt{K} + qA(\alpha) & \text{if } mA(\alpha) < 0 \\ -pA(\alpha) \cdot \sqrt{K} + qA(\alpha) & \text{otherwise} \end{cases}$$

where

$$\begin{cases} mA(\alpha) & = m1A \cdot \big(\tanh(p1A \cdot \alpha) + e1A\big)/\alpha + q1A \\ pA(\alpha) & = m2A \cdot \big(\tanh(p2A \cdot \alpha) + e2A\big)/\alpha + q2A \\ qA(\alpha) & = m3A \cdot \big(\tanh(p3A \cdot \alpha) + e3A\big)/\alpha + q3A \end{cases}$$

Instead the asymptotic value of $L(t)$ is described by

$$D(K,\alpha) = mD(K) \cdot \tanh(qD(K)\,\alpha)$$

i	miA	piA	eiA	qiA
1	$-2.87 \cdot 10^{-2}$	$2.91 \cdot 10^{-1}$	$-2.59 \cdot 10^{-1}$	$6.41 \cdot 10^{-4}$
2	$-7.94 \cdot 10^{-1}$	$2.01 \cdot 10^{-1}$	$-3.02 \cdot 10^{-1}$	$-3.22 \cdot 10^{-2}$
3	4.29	$-1.52 \cdot 10^{-1}$	$2.65 \cdot 10^{-1}$	3.41

Table 2: *Table containing the parameters relative to the expression of $A(K, \alpha)$.*

with

$$\begin{cases} mD(K) & = (N_0 - q1D) \cdot \tanh(p1D \cdot (x - e1D)) + q1D \\ qD(K) & = m2D \cdot \Big(1 - 1/(1 + (e2D/x)^{p2D})\Big) + q2D \end{cases}$$

i	miD	piD	eiD	qiD
1	–	$4.23 \cdot 10^{-4}$	454	848
2	1.75	$8.20 \cdot 10^{-1}$	28.6	$9.75 \cdot 10^{-2}$

Table 3: *Table containing the parameters relative to the expression of $D(K, \alpha)$.*

Finally, the function describing t^*, the instant at which $L(t^*) = (N_0 + D(K, \alpha))/2$, is

$$t^*(K, \alpha) = mT(\alpha) \cdot K - pT(\alpha) \cdot \sqrt{K} + qT(\alpha)$$

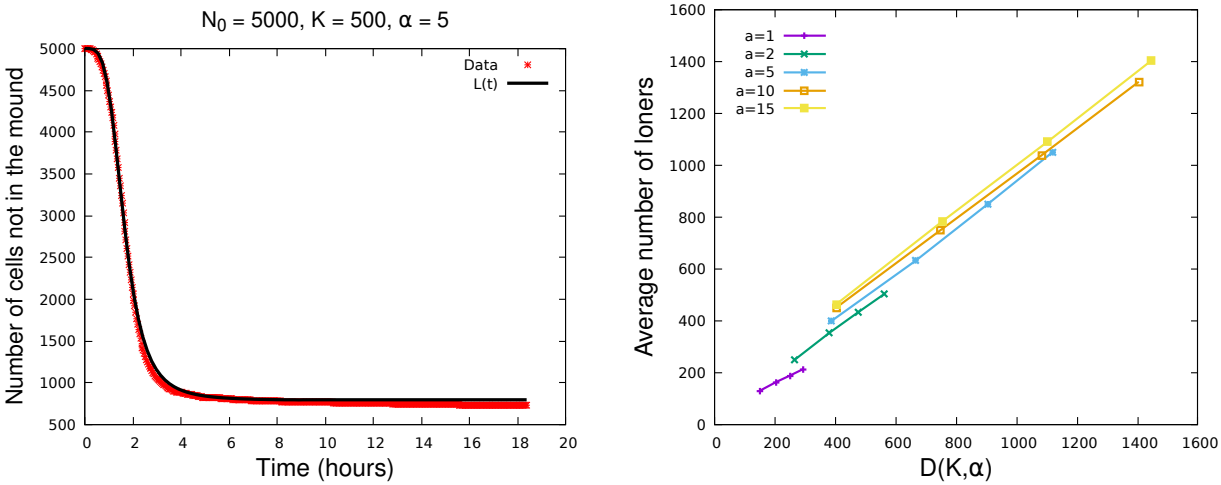
$$\begin{cases} mT(\alpha) & = m1T \cdot \big(\tanh(p1T \cdot \alpha) + e1T\big)/\alpha + q1T \\ pT(\alpha) & = m2T \cdot \big(\tanh(p2T \cdot \alpha) + e2T\big)/\alpha + q2T \\ qT(\alpha) & = m3T \cdot \big(\tanh(p3T \cdot \alpha) + e3T\big)/\alpha + q3T \end{cases}$$

i	miT	piT	eiT	qiT
1	$7.49 \cdot 10^{-3}$	1.5	$-8.20 \cdot 10^{-1}$	$2.15 \cdot 10^{-4}$
2	$6.52 \cdot 10^{-2}$	$6.70 \cdot 10^{-1}$	$-6.11 \cdot 10^{-1}$	$5.17 \cdot 10^{-3}$
3	$2.16 \cdot 10^{-1}$	$5 \cdot 10^{-1}$	$-4.87 \cdot 10^{-1}$	$1.82 \cdot 10^{-1}$

Table 4: *Table containing the parameters relative to the expression of $t^*(K, \alpha)$.*

Two examples of how effective is the empirical function are presented below. The first simply confronts $L(t)$ with the data from the simulations, while in the second the average number of loners for different K's and α's is confronted with the value predicted by $D(K, \alpha)$. The same

expressions proved to be able to describe the data corresponding to different values of N_0, but requiring parameters very different from those presented here.



(a) *Confront between the empirical function and the data.*

(b) *Average number of loners as a function of* $D(K, \alpha)$.

Figure 14: *Plots which exemplify the effectiveness of* $L(t)$.

References

[1] *Eco-evolutionary significance of "loners"*; Rossine FW, Martinez-Garcia R, Sgro AE, Gregor T, Tarnita CE (2020). PLoS Biol 18(3): e3000642. https://doi.org/10.1371/journal.pbio.3000642

[2] *The social amoebae: the biology of cellular slime molds*; Bonner JT.Princeton, NJ: Princeton University Press; 2009.

[3] *The Transforms and Applications Handbook: Second Edition.*; Ed. Alexander D. Poularikas, Boca Raton: CRC Press LLC, 2000

[4] *Local accumulation times for source, diffusion, and degradation models in two and three dimensions*; Peter V. Gordon, Cyrill B. Muratov, and Stanislav Y. Shvartsman. J. Chem. Phys. 138, 104121 (2013); https://doi.org/10.1063/1.4793985

[5] *The Hill equation: a review of its capabilities in pharmacological modelling*; Goutelle S, Maurin M, Rougier F, Barbaut X, Bourguignon L, Ducher M, Maire P. Fundam Clin Pharmacol. 2008 Dec;22(6):633-48. doi: 10.1111/j.1472-8206.2008.00633.x. PMID: 19049668.

[6] *A model for individual and collective cell movement in Dictyostelium discoideum*; Palsson E, Othmer H. Proc Natl Acad Sci U S A. 2000; 97(19):10448–53. https://doi.org/10.1073/pnas.97.19.10448 PMID: 10984537