

Laws of spatially structured population dynamics on a lattice

Natalia L. Komarova¹, Ignacio A. Rodriguez-Brenes¹, and Dominik Wodarz^{2,1}

¹ Department of Mathematics, University of California Irvine, Irvine CA 92617

² Department of Population Health and Disease Prevention, Program in Public Health Susan and Henry Samueli College of Health Sciences, University of California Irvine, Irvine CA 92697

Abstract. We consider spatial population dynamics on a lattice, following a type of a contact (birth-death) stochastic process. We show that simple mathematical approximations for the density of cells can be obtained in a variety of scenarios. In the case of a homogeneous cell population, we derive the cellular density for a 2D spatial lattice with an arbitrary number of neighbors, including the von Neumann, Moore, and hexagonal lattice [1]. We then turn our attention to evolutionary dynamics, where mutant cells of different properties can be generated [2]. For disadvantageous mutants, we derive an approximation for the equilibrium density representing the selection-mutation balance. For neutral and advantageous mutants, we show that simple scaling (power) laws for the numbers of mutants in expanding populations hold in 2D and 3D, under both flat (planar) and range population expansion. These models have relevance for studies in ecology and evolutionary biology, as well as biomedical applications including the dynamics of drug resistant mutants in cancer and bacterial biofilms.

1 Introduction

Much has been said about the existence of laws in biology. Physicists in particular often feel that some beautiful (and hopefully simple) formal relations must exist to help us navigate the complexities of life. Before we begin describing some of our attempts to derive a number of mathematical laws for biological population dynamics, one of us (NK) would like to express her eternal gratitude to Professor Tribelsky, who was her teacher in the Physics Department, Moscow State University in the beginning of 1990s. Without your advice and guidance, I would not be what I am now. Your relentless optimism (often disguised) has taught me to overcome whatever difficulties life posed. I always follow your principle that problems (in life as well as in science) must be addressed at the same pace, and not faster, than they are thrown at you ("Проблемы нужно решать по мере их поступления"). Your "special course" delivered to the theoreticians from the Low Temperature Division, opened up a universe of diverse phenomena that are at the same time fascinating and amenable to understanding. You helped instill the sense of wonder at the beauty of the surrounding world, and this has stayed with me no matter what subject matter I happened

Lattice	Geometry	Density
Von Neumann		$\frac{3-4\delta}{3-\delta}$
Moore		$\frac{7-8\delta}{7-\delta}$
Hexagonal		$\frac{5-6\delta}{5-\delta}$
Mass-action		$1 - \delta$
N_b neighbors		$\frac{N_b-1-N_b\delta}{N_b-1-\delta}$

Table 1: Steady-state densities for different types of lattices, as well as the mass-action system and a generalized, N_b -neighbor system, see [1]. Here $\delta = D/L$, the division-to-death rate ratio.

to focus on, from language evolution to virus dynamics.

In this paper we look for universal biological laws in the area of population dynamics, which is an area relevant for studies in ecology and evolutionary biology. A rich literature exists about the evolutionary dynamics of mutant spread and invasion, investigating aspects such as fixation probabilities and fixation times of different kinds of mutants in constant populations [3, 4] as well as mutant dynamics in growing populations [5, 6, 7], where mathematical approaches were motivated by the famous Luria-Delbruck experiments [8]. Traditionally, much of this work has focused on well-mixed (non-spatial), homogeneous populations. In a wide array of biological scenarios, however, cells and organisms evolve in more complex settings, such as in spatially structured habitats (bacterial biofilms, cells in tissues and tumors), and within heterogeneous population structures (such as stem and more differentiated cells in tissues [9] or bacteria with different degrees of specialization in biofilms [10, 11]). In the last 15 years, theoretical work has extended our understanding of mutant emergence in spatially structured populations [12, 13, 14, 15, 16, 17, 18, 19, 20]. An excess of mutational jackpot events was observed in spatial compared to well-mixed systems; such events result from mutations arising at the surface of expanding, spatially structured populations, surfing at the edge of range expansions, and appearing as mutant “bubbles” or “slices”. These jackpot events have been implicated in the finding that the average number of neutral mutants when the total population reaches a given threshold size is significantly larger in spatial compared to non-spatial settings [20]. Further work by our group [2] showed that the evolutionary dynamics of mutants in spatially structured, expanding populations are characterized by additional complexities.

In general, spatial population dynamics are significantly more complex com-

	2D flat	2D range	3D flat	3D range	Exponential
Mutant property					
Disadvantageous	uN	uN	uN	uN	uN
Neutral	uN^2	$uN^{3/2}$	uN^2	$uN^{4/3}$	$uN \ln N$
Advantageous	uN^3	uN^2	uN^4	uN^2	$uN^{(2\alpha-1)/\alpha}$

Table 2: Mutant scaling laws for expanding populations, see [2]. Here N is the population size, u is the mutation rate, and $\alpha > 1$.

pared to the mass-action dynamics, and analytical approximations are not easily derived. In this paper we describe two examples where our group has been successful in obtaining analytical approximations for laws of spatial dynamics: (a) Steady state density in space; and (b) Scaling laws for the number of mutants in expanding populations.

(a) Analytical approximations for steady-state density. In [1], we considered the phenomenon of range expansion, the process in biology by which a species spreads to new areas. We derived numerical methodology that allowed for efficient computations of the number of individuals as the species expands its range in space. In addition, we found approximations for the steady-state density of populations, or the core density of expanding colonies, in several cases, such as the von Neumann and Moore neighborhoods on a square lattice, and for a honeycomb neighborhood on a hexagonal lattice, see table 1. For the same death-to-birth ratio, the grid is more packed under the Moore neighborhood, because of the availability of more neighbors per site. As a consequence, the equilibrium density corresponding to the Moore neighborhood is higher and closer to that of mass action. A general formula that provides an approximation for the steady-state density as a function of the number of neighbors is also given (table 1). In this paper we provide a novel mathematical derivation of these laws, which generalizes to any neighborhood size (section 3).

(b) Laws of disadvantageous, neutral, and advantageous mutant spread in different geometries. In [2], we studied the evolutionary dynamics of disadvantageous, neutral, and advantageous mutants in spatially expanding populations, where the growing front was characterized by different symmetries in 2D and 3D, see table 2. In particular, while disadvantageous mutants grow linearly with N (and differences among different cases are more subtle), neutral mutants grow as a decreasing power of N as we go from 2D flat front (which is essentially a 1D growth), to 2D and 3D range. There are always fewer neutral or advantageous mutants in the mass action (exponentially growing) case compared to any spatially restricted growth. Here (Sections 4 and 5) we provide a derivation of these laws, which follows [2].

In the case of disadvantageous mutants, the number of mutants always scales

with the 1st power of N . In addition to this scaling law, for the case of disadvantageous mutants, we also derive an approximate expression for the number of mutants in quasi steady state, given that the number of wild type cells is N :

$$\# \text{ mutants} = \frac{D_w L_w (4D_m^2 + 3D_m L_w + D_w L_m) u N}{D_m (4D_m + 3L_w) (D_m L_w - D_w L_m)}, \quad (1)$$

see Section 4.¹ We show that the proportion of mutants at selection-mutation balance is higher for spatially distributed systems compared to well-mixed systems at equilibrium.

2 General set-up and agent-based modeling of population dynamics

In order to describe spatial growth and turnover of cells, as well as population dynamics of species, we will use a continuous time Markov process which is a generalization of the usual birth-death process and a type of a contact process. We will assume that individuals of one or two different types exist, which we will call "wild type" and "mutant" individuals.² At the core of the description is a lattice that specifies possible locations of cells. This can be viewed as a geometric network, where each node is connected with its neighbors. The state space consists of different locations of the cells of the two types on the lattice. Cells are characterized by division and death rate parameters, which are denoted by L_w and D_w for division and death rates of the wild-type cells, and L_m and D_m for division and death rates of mutants. In the case where only one type of cells is considered, the notation is simply L and D for the two rates.

Each of the nodes can either be empty or contain one cell of either type. During an infinitesimally small time-increment, Δt , a given wild-type cell can attempt a division with probability L_w and death with probability D_w . If division is attempted, then an offspring is placed in one of the neighboring locations, chosen randomly and uniformly; division only happens if the chosen node is empty. The offspring cell is wild type with probability $1 - u$ and mutant with probability u , where $0 \leq u < 1$ is the mutation rate. Mutant cells divide and die according to similar rules, except in the models considered here we always assume that an offspring of a mutant cell is a mutant cell.

A number of biological phenomena can be studied by slight modifications of this process. To study the growth laws in the absence [1] or presence [2] of mutants we assume an infinitely large grid. To study quasi-steady state (which describes e.g. the turnover of cells in homeostasis) we make the grid finite and impose relevant boundary conditions.

For numerical explorations of these processes, agent-based modeling (ABM) is often used, see e.g. [21, 22, 23]. For example, consider a 2-dimensional ABM

¹This expression is valid for a 2D von Neumann grid, and can be generalized to other cases.

²To reflect the versatility of these models we will use both "individuals" and "cells" to refer to the biological agents under consideration.

on a square grid with individuals of two types. A spot on the grid can be empty or can contain a cell, which is either wild-type or mutant. At each time step, the grid is randomly sampled N times, where N is the total number of cells currently in the system. If the sampled cell is wild-type, the cell attempts division (described below) with a probability proportional to L_w or dies with a probability proportional to D_w . When reproduction is attempted, a target spot is chosen randomly among the N_b nearest neighbors of the cell. A neighborhood may contain e.g. $N_b = 4$ cells (the Von Neumann neighborhood) or $N_b = 8$ neighbors (the Moore neighborhood). If that spot is empty, the offspring cell is placed there. If it is already filled, the division event is aborted.³ The offspring cell is assigned wild type with probability $1 - u$ and it is a mutant with probability u . If the sampled spot contains a mutant cell, the same processes occur. Attempted division occurs with a probability $\propto L_m$, and the cell dies with a probability $\propto D_m$. Initial and boundary conditions are determined by the specific geometric setting investigated. For 2D spatial simulations, an $n \times n$ square, or an $n \times w$ rectangular domain could be considered. At the boundaries of the domain, a spot is assumed to have fewer neighbors, i.e more division events will fail. The process starts with a single wild-type cell, placed e.g. into the center of the grid. Simulations always stop before the boundary of the grid is reached.

In what follows below we derive some approximations of important observables from these types of dynamics, which have a clear biological meaning. We will discuss both single-type populations and the co-dynamics of wild-type and mutant individuals.

3 Analytical approximations for steady-state density (a single type)

In this section we consider populations consisting of only a single cell type with division and death rates L and $D < L$, and no mutations ($u = 0$).

3.1 Equilibrium density in space: an analytical method

The process of range expansion (colonization) is one of the basic types of biological dynamics, whereby a species grows and spreads outwards, occupying new territories. Spatial modeling of this process is naturally implemented as a stochastic ABM of the type described above, with individuals (in this case, of only a single type) occupying nodes on a rectangular grid, births and deaths occurring probabilistically, and individuals only reproducing onto un-occupied neighboring spots. This approach is known to be computationally expensive. In [1], we derived a set of efficient computational tools which showed to be in good agreement with the underlying stochastic process of spatial expansion. As part

³This modeling choice represents the assumption that the probability of divisions is reduced under more crowded conditions. This is similar to the logistic growth term often used in deterministic models.

of the method development, we were able to obtain approximate expressions for the quasi-steady state (core) density of the individuals for different types of grid. In [1], we provided the density formula for the contact process,

$$\rho = \frac{N_b - 1 - N_b \delta}{N_b - 1 - \delta}, \quad (2)$$

where $\delta = D/L$ and N_b is the number of neighbors. This formula was derived for several cases of N_b , but no general derivation that would work for a given N_b was supplied. Here we present such a derivation, starting with $N_b = 2$ and generalizing to any N_b .

In 1D, with only $N_b = 2$ neighbors, equation (2) gives

$$\rho = \frac{1 - 2\delta}{1 - \delta}, \quad (3)$$

which, although a slight overestimation of the density, provides a good approximation. To derive this formula, consider a 1D ABM of the type described above. At the equilibrium, denote by p_1 the probability that a cell is located next to an existing cell, and p_0 the probability that a cell is located next to an existing empty spot. These two quantities can be estimated numerically.

A 1D realization of the population at a fixed moment of time can be viewed as a Markov chain with states $\{0, 1\}$, where state 0 corresponds to the absence of a cell at a given spot and state 1 denotes the presence of a cell at a given spot. The transition probability matrix is given by

$$P = \begin{pmatrix} p_1 & 1 - p_1 \\ p_0 & 1 - p_0 \end{pmatrix}$$

where P_{11} is the probability that there is a cell on the right of a cell, P_{12} is the probability that the spot on the right of a cell is empty, P_{21} is the probability that there is a cell on the right of an empty spot, and P_{22} is the probability that an empty spot is on the right of an empty spot. The steady state probability distribution of this process is $(\rho, 1 - \rho)$, where ρ is the probability that a given spot contains a cell. This can be found as the eigenvector corresponding to the unit eigenvalue, and is given by

$$\rho = \frac{p_0}{1 - p_1 + p_0}. \quad (4)$$

This quantity (given the numerically calculated p_0 and p_1) is a very good approximation of the actual (numerical) density. Another way to derive this connection between ρ and p_0, p_1 is as follows:

$$\rho = \rho p_1 + (1 - \rho) p_0, \quad (5)$$

where the left hand side is the probability to have a cell at a given point, the first term of the right assumes that there is a cell to the left of that point (ρ), and then the given point contains a cell with probability p_1 ; and the second

term of the right assumes that there is no cell to the left of that point $(1 - \rho)$, and then the given point contains a cell with probability p_0 . Solving this for ρ gives expression (4).

Now, suppose that a cell is located at a given location. Then a cell is located to the right of it (which we will call the focal location) with probability p_1 . We have, after a single update of the contact process:

$$p_1 = (1 - p_1) \left(\frac{L}{2} + p_0 \frac{L}{2} \right) \frac{1}{N} + p_1 \left(1 - \frac{D}{N} \right). \quad (6)$$

The first term on the right assumes that there was no cell at the focal location $(1 - p_1)$ but a cell on its left was chosen (probability $1/N$) that divided to its right $(L/2)$ or that a cell on its right exists (probability p_0), was chosen $(1/N)$ and divided to its left $(L/2)$. The second term assumes that there was a cell at the focal location (p_1) and it did not die $(1 - D/N)$. The equality follows from the assumption of having an equilibrium.

Similarly, we can assume that there is no cell at a given location, then to its right (the focal location) the probability to have a cell, p_0 , satisfies:

$$p_0 = (1 - p_0) p_0 \frac{L}{2N} + p_0 \left(1 - \frac{D}{N} \right). \quad (7)$$

The first term on the right assumes that there was no cell at the focal location $(1 - p_0)$ but a cell on its right exists (probability p_0), was chosen $(1/N)$ and divided to its left $(L/2)$. The second term assumes that there was a cell at the focal location (p_0) and it did not die $(1 - D/N)$.

Solving equations (6) and (7) for p_0 and p_1 , we obtain:

$$p_1 = 1 - \frac{D}{L}, \quad p_0 = 1 - \frac{2D}{L}.$$

Substituting this into formula (4), we obtain expression (3).

This analysis easily generalizes to other systems (including higher dimensionalities), where the number of neighbors is given by N_b . Instead of equations (6) and (7), we have

$$p_1 = (1 - p_1) \left(\frac{L}{N_b} + \frac{(N_b - 1)p_0 L}{n} \right) \frac{1}{N} + p_1 \left(1 - \frac{D}{N} \right), \quad (8)$$

$$p_0 = (1 - p_0) \frac{(N_b - 1)p_0 L}{N_b N} + p_0 \left(1 - \frac{D}{N} \right). \quad (9)$$

Solving this system, we obtain

$$p_1 = 1 - \frac{D}{L}, \quad p_0 = 1 - \frac{N_b D}{(N_b - 1)L}.$$

Substitution into formula (4) (which holds in these systems because equation (5) remains valid in these systems), yields formula (2).

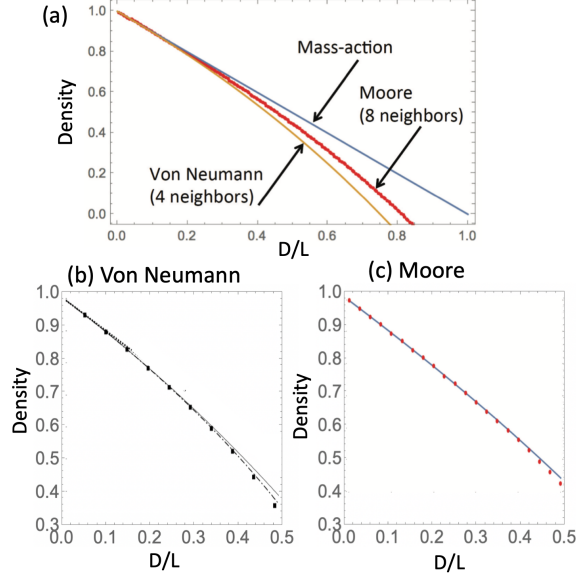


Figure 1: Quasi-equilibrium density of cells. (a) Equation (2), showing the equilibrium density as a function of D/L for $N_b = 4$ (von Neumann lattice), $N_b = 8$ (Moore lattice), and $N_b \rightarrow \infty$ (mass-action). (b) Comparison of formula (2) (solid line) with numerically calculated density values (dots) for $N_b = 4$. Dash-dot line represents a higher-order approximation, see [1] for details. (c) Comparison of formula (2) (solid line) with numerically calculated density values (dots) for $N_b = 8$.

Figure 1 plots the quasi-equilibrium density approximated by formula (2) as a function of D/L for two finite values of N_b corresponding to the von Neumann and Moore neighborhoods. Panel (a) compares them with the well-known mass-action result, showing that higher connectivity of the underlying network corresponds to a higher density of the cells. Panels (b) and (c) compare the results with numerical simulations, demonstrating that formula (2) provides a very good approximation. We notice that the approximation works better at higher densities, compared to lower ones. In the derivation we assumed that the state of a given position on the lattice only depends on its immediate neighbor, and not on other, more distant, sites. This assumption becomes less valid at low densities, because macroscopic structures with strong correlations over distances > 1 form. Therefore, the approximation becomes worse as the death rate approaches the division rate.

4 Disadvantageous mutants and selection-mutation balance in spatial models

In the remainder of the paper we consider two types of individuals, wild type and mutants. This section deals with disadvantageous mutants (defined below) while section 5 focuses on neutral and advantageous mutants.

4.1 A basic ODE formulation

Let us denote the wild type population as $x(t)$ and the mutant population as $y(t)$. Suppose that mutations happen at the rate u , and the division and death rates of wild type cells are L_w and $D_w < L_w$, while the division and death rates of mutants are L_m and D_m . Then the competition dynamics of cells in a mass-action system can be formulated as follows:

$$\dot{x} = L_w x(1-u) \left(1 - \frac{x+y}{\mathcal{K}}\right) - D_w x, \quad (10)$$

$$\dot{y} = (L_w x u + L_m y) \left(1 - \frac{x+y}{\mathcal{K}}\right) - D_m y, \quad (11)$$

where $\mathcal{K}$ is the carrying capacity. This system illustrates differences in the behavior of advantageous and disadvantageous mutants. If

$$\frac{D_m}{D_w} - \frac{L_m}{L_w(1-u)} < 0,$$

then the globally stable solution is $x = 0$, $y = \mathcal{K}(1 - D_m/L_m)$, that is, the mutant excludes the wild type and takes over. In this case we say that the mutant is advantageous. If the inequality above is reversed, the mutants are disadvantageous and the so-called selection-mutation balance is reached, where mutants and wild type cells coexist. This happens because the (stronger) wild type cells produce mutants at a nonzero rate u , but the latter cannot take over and remain at a fraction of the population. For simplicity let us assume that the mutation rate is low:

$$u \ll \frac{D_m}{D_w} - \frac{L_m}{L_w} \equiv \gamma.$$

Then we can say that the mutants are disadvantageous if

$$\frac{D_m}{D_w} - \frac{L_m}{L_w} > 0, \quad (12)$$

and then the equilibrium solution is given by

$$\bar{x} = \mathcal{K} \left(1 - \frac{D_w}{L_w}\right), \quad (13)$$

where we neglected terms of the order u/γ , and the number mutants is

$$\bar{y} = \bar{x}u \left(\frac{D_m}{D_w} - \frac{L_m}{L_w} \right)^{-1}, \quad (14)$$

where we neglected terms of the order $(u/\gamma)^2$.

Below we calculate equilibrium densities of disadvantageous mutants and (advantageous) wild type cells in a spatially distributed system at steady state. This will also correspond to the densities in the core of an expanding system away from the advancing front.

4.2 A spatial description: equations for the densities

We restrict our description to a 2D square grid, with the von Neumann neighborhood (that is, each location has 4 nearest neighbors); the methodology is generalizable to the Moore neighborhood (8 neighbors). We use a method similar to that of [1]. Two random variables describe the state of the stochastic system at each spatial location, x : ρ_x describes wild type cells, such that

$$\rho_x = \begin{cases} 1, & \text{if a wild type cell is at location } x, \\ 0, & \text{otherwise,} \end{cases}$$

and η_x describes mutant cells, such that

$$\eta_x = \begin{cases} 1, & \text{if a mutant cell is at location } x, \\ 0, & \text{otherwise.} \end{cases}$$

Note that ρ_x and η_x cannot be equal to one simultaneously; an empty spot corresponds to $\rho_x = \eta_x = 0$. We assume that wild type cells have division and death rates L_w and D_w , and mutant cells have division and death rates L_m and D_m . Wild type cells mutate with probability u , and no back mutations are considered.

Denote the expectation of ρ_x and η_x by

$$\langle \rho_x \rangle = \rho, \quad \langle \eta_x \rangle = \eta,$$

where we assumed that the expected values do not depend on spatial location, since we are interested in spatially homogeneous equilibrium solutions. We have

$$\dot{\rho} = \left\langle \frac{L_w}{N_b} (1-u)(1-\rho_x)(1-\eta_x) \sum_k \rho_x^{(k)} - D_w \rho_x \right\rangle, \quad (15)$$

$$\dot{\eta} = \left\langle \frac{1}{N_b} (1-\rho_x)(1-\eta_x) \sum_k \left(L_w u \rho_x^{(k)} + L_m \eta_x^{(k)} \right) - D_m \eta_x \right\rangle, \quad (16)$$

where the product $(1-\rho_x)(1-\eta_x)$ is nonzero only if location x is empty, and the summation goes over all the neighbors of point x , which reproduce into location x at rates L_w/N_b and L_m/N_b if they are wild type or mutant, respectively. The

superscript in the notation $\rho_x^{(k)}$, $\eta_x^{(k)}$ refers to the quantity at a location, k , neighboring the focal location, x .

Let us consider the von Neumann neighborhood ($N_b = 4$). In the right hand side of equation (15), the terms in the summation have the form

$$\langle (1-\rho_x)(1-\eta_x)\rho_x^{(k)} \rangle = \langle \rho_x^{(k)} \rangle - \langle \rho_x \rho_x^{(k)} \rangle - \langle \rho_x^{(k)} \eta_x \rangle + \langle \rho_x^{(k)} \rho_x \eta_x \rangle = \rho - W - I, \quad (17)$$

and in equation (16) there are also terms of the form

$$\langle (1-\rho_x)(1-\eta_x)\eta_x^{(k)} \rangle = \langle \eta_x^{(k)} \rangle - \langle \rho_x \eta_x^{(k)} \rangle - \langle \eta_x^{(k)} \eta_x \rangle + \langle \eta_x^{(k)} \rho_x \eta_x \rangle = \eta - I - M. \quad (18)$$

In the expressions above, we have $\langle \rho_x \rho_x^{(k)} \eta_x^{(k)} \rangle = 0$, because either $\eta_x^{(k)}$ or $\rho_x^{(k)}$ is zero at location $x^{(k)}$, and the three types of dyads are defined as follows:

- $W = \langle \rho_x \rho_x^{(k)} \rangle$ is the probability to have two wild type cells at two neighboring locations,
- $I = \langle \rho_x \eta_x^{(k)} \rangle$ is the probability to have a wild type cell and a mutant at two neighboring locations,
- $M = \langle \eta_x \eta_x^{(k)} \rangle$ is the probability to have two mutant cells at two neighboring locations.

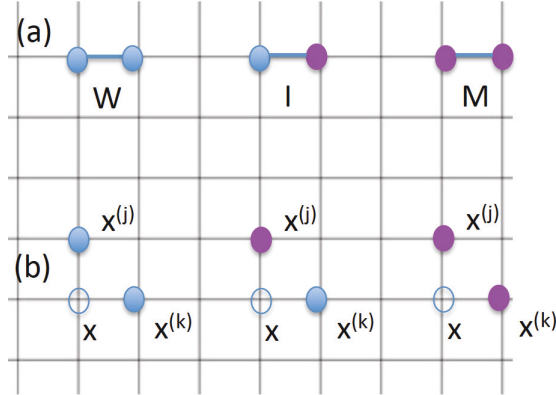


Figure 2: Steps in the derivation of equations for a two-component system of wild type and mutant cells. Blue circles denote wild type, and purple denote mutant cells. (a) Three configurations, whose correlations appear in equations (19) and (20). (b) Three types of correlations needed for equations for W , I , and M .

Figure 2(a) illustrates these three configurations. In terms of these three correlations, equations (15) and (16) can be rewritten as

$$\dot{\rho} = L_w(1-u)(\rho - W - I) - D_w\rho, \quad (19)$$

$$\dot{\eta} = L_w u(\rho - W - I) + L_m(\eta - I - M) - D_m\eta. \quad (20)$$

The correlations for the three dyads that appear in these equations require their own equations to close the system. Let us derive an equation for W . We have

$$\dot{W} = \left\langle 2(1 - \rho_x)(1 - \eta_x)\rho_x^{(k)} \sum_j \rho_x^{(j)} \frac{L_w}{N_b}(1 - u) - 2D_w W \right\rangle,$$

where we assume that one of the points in the dyad contains a wild type cell (term $\rho_x^{(k)}$), while the other point is empty (term $(1 - \rho_x)(1 - \eta_x)$), and that one of its neighbors (location $x^{(j)}$) contains a wild type cell, which reproduces faithfully into point x at rate $L_w(1 - u)/N_b$. Note that either of the two points could be empty, which results in the multiplier 2 in the first term on the right hand side. Similarly, either of the dyad's locations can experience cell death, resulting in the negative rate $2D_w$. In order to calculate the average, we need to consider terms

$$\langle (1 - \rho_x)(1 - \eta_x)\rho_x^{(k)}\rho_x^{(j)} \rangle. \quad (21)$$

Note that here and below, the operation of averaging makes the expression independent on the actual location x . Further, the superscripts (k) and (j) do not refer to any specific neighbor of x , but to any neighbor of x ; in particular, location $x^{(j)}$ may be the same or different than location $x^{(k)}$. In the case when the two locations are different, correlation (21) is presented in figure 2(b), on the left.

In equations for $\dot{M}$ and $\dot{I}$, the following expressions appear in addition to (21):

$$\langle (1 - \rho_x)(1 - \eta_x)\rho_x^{(k)}\eta_x^{(j)} \rangle, \quad \langle (1 - \rho_x)(1 - \eta_x)\eta_x^{(k)}\eta_x^{(j)} \rangle.$$

These correlations are shown in figure 2(b), center and right. Therefore, denoting by a and b either ρ or η , we evaluate the average of the form

$$\langle (1 - \rho_x)(1 - \eta_x)a_x^{(k)}b_x^{(j)} \rangle, \quad (22)$$

which corresponds to a dyad with one of the locations (location x) empty, and the other (location $x^{(k)}$) containing type "a", while a neighbor of x (location $x^{(j)}$) contains type "b". First let us assume that location $x^{(j)}$ is different from location $x^{(k)}$. Under von Neumann neighborhoods this implies that $x^{(j)}$ are $x^{(k)}$ are not each other's neighbors, because on a square grid, there could not be a non-degenerate triangle with diameter 1 or less. Expression (22) is equal to

$$\begin{aligned} & P(b_x^{(j)} = 1 | a_x^{(k)} = 1, \rho_x = \eta_x = 0) P(a_x^{(k)} = 1, \rho_x = \eta_x = 0) \approx \\ & P(b_x^{(j)} = 1 | \rho_x = \eta_x = 0) P(a_x^{(k)} = 1, \rho_x = \eta_x = 0) = \\ & \frac{P(b_x^{(j)} = 1, \rho_x = \eta_x = 0) P(a_x^{(k)} = 1, \rho_x = \eta_x = 0)}{P(\rho_x = \eta_x = 0)} \end{aligned} \quad (23)$$

The expression in the denominator is calculated as follows:

$$P(\rho_x = \eta_x = 0) = \langle (1 - \rho_x)(1 - \eta_x) \rangle = \langle 1 - \rho_x - \eta_x + \rho_x\eta_x \rangle = 1 - \rho - \eta.$$

Depending on the types at location x , the expressions in the numerator of (23) can be of two types:

$$P(\rho_x^{(k)} = 1, \rho_x = \eta_x = 0) \text{ or } P(\eta_x^{(k)} = 1, \rho_x = \eta_x = 0),$$

and they are calculated in (17) and (18) respectively.

Next, we assume that location $x^{(j)}$ is the same as $x^{(k)}$. Then, if types "a" and "b" in expression (22) are different, then we obtain $\langle (1 - \rho_x)(1 - \eta_x)\rho_x^{(k)}\eta_x^{(k)} \rangle = 0$. If the types are the same, then we obtain expression (17) or (18). To summarize, expressions of type (22) are given as follows:

$$\begin{aligned} \langle (1 - \rho_x)(1 - \eta_x)\rho_x^{(k)}\rho_x^{(j)} \rangle &= \begin{cases} \frac{(\rho - W - I)^2}{1 - \rho - \eta}, & x^{(j)} \neq x^{(k)}, \\ \rho - W - I, & x^{(j)} = x^{(k)}, \end{cases} \\ \langle (1 - \rho_x)(1 - \eta_x)\rho_x^{(k)}\eta_x^{(j)} \rangle &= \begin{cases} \frac{(\rho - W - I)(\eta - I - M)}{1 - \rho - \eta}, & x^{(j)} \neq x^{(k)}, \\ 0, & x^{(j)} = x^{(k)}, \end{cases} \\ \langle (1 - \rho_x)(1 - \eta_x)\eta_x^{(k)}\eta_x^{(j)} \rangle &= \begin{cases} \frac{(\eta - I - M)^2}{1 - \rho - \eta}, & x^{(j)} \neq x^{(k)}, \\ \eta - I - M, & x^{(j)} = x^{(k)}. \end{cases} \end{aligned}$$

The equation for W is then given by

$$\dot{W} = \frac{L_w}{2}(1 - u) \left(\rho - W - I + \frac{3(\rho - W - I)^2}{1 - \rho - \eta} \right) - 2D_w W. \quad (24)$$

Similarly, the other two equation can be derived:

$$\begin{aligned} \dot{I} &= \frac{3}{4}[L_w(1 - u) + L_m] \frac{(\rho - W - I)(\eta - I - M)}{1 - \rho - \eta} + \frac{L_w u}{4} L_m \left(\rho - W - I + \frac{3(\rho - W - I)^2}{1 - \rho - \eta} \right) \\ &\quad - (D_w + D_m)I, \end{aligned} \quad (25)$$

$$\dot{M} = \frac{3L_w u}{2} \frac{(\rho - W - I)(\eta - I - M)}{1 - \rho - \eta} + \frac{L_m}{2} \left(\eta - I - M + \frac{3(\eta - I - M)^2}{1 - \rho - \eta} \right) - 2D_m M. \quad (26)$$

The closed system of equations for ρ, η, W, I , and M is given by equations (19), (20), (24), (25), and (26).

4.3 Selection mutation balance solution

Solving these equations in steady state exactly is difficult, but if the mutation rate $u \ll 1$, we can find the approximate solution. We start by setting $u = 0$ and obtaining the steady state solution. Apart from the trivial solution and a negative solution, there are two symmetric solutions where only one type survives (competitive exclusion). We will use the one where the wild type excludes mutants:

$$\rho^{(0)} = 1 + \frac{3D_w}{D_w - 3L_w}, \quad W^{(0)} = 1 - \frac{6D_w}{D_w - 3L_w} - \frac{4D_w}{L_w}, \quad \eta^{(0)} = I^{(0)} = M^{(0)} = 0, \quad (27)$$

where the superscript corresponds to the zeroth order in the expansion in terms of small u . Note that as expected, the expression for $\rho^{(0)}$ coincides with the one given in table 1 for von Neumann grid ($N_b = 4$). We then look for the first correction by substituting

$$\rho = \rho^{(0)} + u\rho^{(1)}, \quad \eta = u\eta^{(1)}, \quad W = W^{(0)} + uW^{(1)}, \quad I = uI^{(1)}, \quad M = uM^{(1)},$$

inserting in the system of 5 equations, keeping only the first order of expansion in u , and solving for $\rho^{(1)}, \dots, M^{(1)}$. We obtain

$$\eta = \eta^{(1)}u = \frac{D_w L_w (4D_w - 3L_w)(4D_m^2 + 3D_m L_w + D_w L_m)u}{D_m (D_w - 3L_w)(4D_m + 3L_w)(D_m L_w - D_w L_m)}. \quad (28)$$

This is the equilibrium solution corresponding to mutation-selection balance in the presence of spatial interactions. This approach is valid as long as the wild type is advantageous (inequality (12)). In the opposite scenario, this solution is unstable, and the system converges to the mutants excluding the wild type.

Under selection-mutation balance, of interest is the equilibrium proportion of mutants in the system given by

$$\nu_{vN}^{eq} = \frac{\eta^{(1)}u}{\rho^{(0)}} = \frac{D_w L_w (4D_m^2 + 3D_m L_w + D_w L_m)u}{D_m (4D_m + 3L_w)(D_m L_w - D_w L_m)}. \quad (29)$$

It is instructive to compare this quantity with the equilibrium proportion of mutants in a mass-action system, $\nu_{m-a}^{eq} = \bar{y}/\bar{x}$, equation (14):

$$\nu_{m-a}^{eq} = \frac{D_w L_w u}{D_m L_w - D_w L_m}.$$

We have

$$\frac{\nu_{vN}^{eq}}{\nu_{m-a}^{eq}} = 1 + \frac{D_w L_w}{D_m (4D_m + 3L_w)} > 1.$$

In other words, the relative contents of mutants (in proportion to the wild-types) are higher in spatially distributed systems compared to mass-action systems at equilibrium. This result resembles our recent analysis presented in [24], where we show that the mean number of disadvantageous mutants is higher in fragmented populations compared to well-mixed populations. A comparison with non-equilibrium, growing well-mixed populations is presented in [2] and not discussed here.

4.4 Applications and comparison with computations

The result in (29) is applicable for two relevant scenarios.

- **Quasi-equilibrium density in finite populations.** If simulations are continued until a finite grid is filled, the population reaches a quasi-equilibrium state where wild type and disadvantageous mutant cells coexist. For a 2D square grid under von Neumann neighborhood, the density

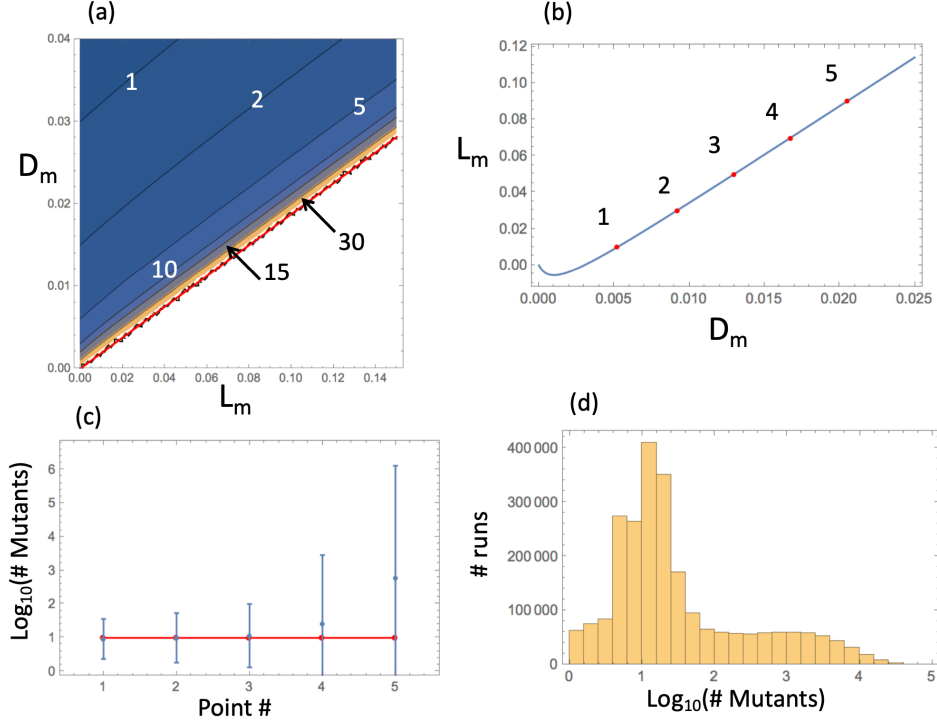


Figure 3: The level of mutants in the spatial (von Neumann) system: analytical approximation and numerical results. (a) The quantity $Nu\nu_{vN}^{eq}$ is presented as a contour plot as a function of L_m and D_m , for fixed values of L_w and D_w . Mutants are disadvantageous above the red line (inequality (12)). The contours' values are specified. (b) Solution L_m of equation $Nu\nu_{vN}^{eq} = 10$ as a function of D_m ; the 5 points used in panel (c) are marked in red and numbered. (c) The comparison of predicted (10, horizontal red line) and simulated number of mutants in the 5 parameter pairs from panel (b). Simulated means and standard deviations are shown (out of 2.5×10^6 runs). (d) For the 5th parameter combination, the numerically obtained histogram of the number of mutants is shown. The rest of the parameters are $u = 2 \times 10^{-5}$, $N = 10^5$, $L_w = 0.08$, and $D_w = 0.015$.

of the mutants is approximated by η in formula (28), while the density of wild types is given by $\rho^{(0)}$, equation (27). The total numbers of mutant and wild type cells are obtained by multiplying the quantities η and $\rho^{(0)}$ by the total number of grid points, respectively. Note that this scenario is not interesting in the case of advantageous or neutral mutants, as the entire population will eventually consist of mutant cells.

- **Number of mutants in spatially expanding populations.** Simulating a growing population (on a large grid where the boundaries are not reached), we can ask how the number of disadvantageous mutants scales with the total number of cells, N . Since the core of the expanding colony is in quasi-equilibrium, equation (29) shows that the number of mutants grows as the first power of N . Results for the scaling laws for neutral and advantageous mutants are discussed in the next section.

The expected number of mutants predicted theoretically was compared with results of numerical simulations. This was done in the following way. At size N , the number of mutants (in the von Neumann case) is predicted to be $Nu\nu_{vN}^{eq}$, see equation (29). Solving the equation $Nu\nu_{vN}^{eq} = const$, we can obtain the pairs (L_m, D_m) of mutant kinetic rates corresponding to a predicted given number of mutants in a system of size N . Figure 3(a) shows the predicted number of mutant as a contour plot. The closer to the “neutrality” line (see inequality (12)) the larger the predicted number of mutants. Solution of equation $Nu\nu_{vN}^{eq} = const$ is shown in panel (b), and for 5 points from the solution set, the predicted number of mutants (given by 10) is compared with the numerically obtained mean (plotted together with the standard deviation, panel (c)). We can see that for larger mutant division rates, the deviation from the theory becomes significant. Panel (d) shows a histogram of numbers of mutants for the parameters corresponding the 5th point. One can see that the distribution has a long tail and a very large standard deviation. This is the consequence of macroscopic structures (“slices”) that cannot be handled by the present, local, method.

5 Laws of neutral and advantageous mutant spread in different geometries

In the previous section we showed that (under the assumption of small mutation rates) the number of disadvantageous mutants in a spatially growing population scales with the total number of cells, as $\propto Nu$. If cells are advantageous or disadvantageous, they obey different scaling laws [2]. In this section we present simple derivations for those laws in different geometries.

5.1 Derivation of the growth laws

Below we consider several scenarios that differ by the dimensionality of the grid (2D or 3D) and by the direction of growth (planar growth vs range expansion, see table 2).

2D flat front. Let us first assume that the death rate of cells is equal to zero. Consider cells growing along the surface of a cylinder of width W . This represents a one-directional growth process, where during each generation, we assume that W new cells appear, and the total population is given by $N =$

$\mathcal{L}W$, where $\mathcal{L}$ represents the number of layers. The value of $\mathcal{L}$ is proportional to the number of generations, and thus to the physical time, t :

$$\mathcal{L} \propto t.$$

The following calculation estimates the growth law of mutants. Every time a new layer (of width W) is added, the mean number of new mutations is given by Wu . Suppose that mutants are neutral. Then, each such mutation will give rise to an array of daughter mutant cells of width 1, see figure 4. The length of this array is given by $\mathcal{L} - i$, where i is the layer at which the mutation occurred. Therefore, the total expected number of neutral mutants is a cylinder of length $\mathcal{L}$ is given by

$$M_{2D \text{ flat}}^{neut} = \sum_{i=1}^{\mathcal{L}} Wu \sum_{j=i+1}^{\mathcal{L}} 1 = \frac{uW\mathcal{L}(\mathcal{L}-1)}{2} \approx \frac{uW\mathcal{L}^2}{2} = \frac{u}{2W}N^2, \quad (30)$$

where we assumed $\mathcal{L} \gg 1$. Note that in this derivation we assumed that the number of mutants is small compared to the total population, and individual mutant clones do not interact. In a more precise calculation, the number of wild type cells in each layer is smaller than W because of the existence of mutants, and thus the rate of new mutant production is smaller than Wu . We, however, assume that $u\mathcal{L}W \ll 1$, such that the number of mutants is relatively small.

Note that the number of neutral mutants decreases with W , see figure 5; the largest number of mutants is achieved in the case of $W = 1$, a one-dimensional expanding array of cells.

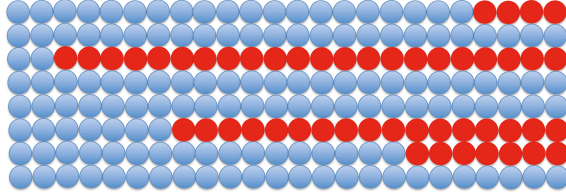


Figure 4: The conceptual model for mutant number calculations, the case of neutral mutants in a colony growing along the surface of a cylinder (2D flat front).

Next, let us consider advantageous mutants. In this case, each new mutant gives rise to a triangular clone. In the first layer, the width of the clone is 1, in the next layer it is $1 + s$, and in the k th layer it is $1 + (k-1)s$, where parameter $s \geq 0$ measures the advantage of the mutants (with $s = 0$ corresponding to neutral mutants). Therefore, we have

$$M_{2D \text{ flat}}^{adv} = \sum_{i=1}^L Wu \sum_{j=i+1}^{\mathcal{L}} (1 + (j - (i+1))s) = uW\mathcal{L}(\mathcal{L}-1) \left(\frac{1}{2} + \frac{s(\mathcal{L}-2)}{6} \right) \quad (31)$$

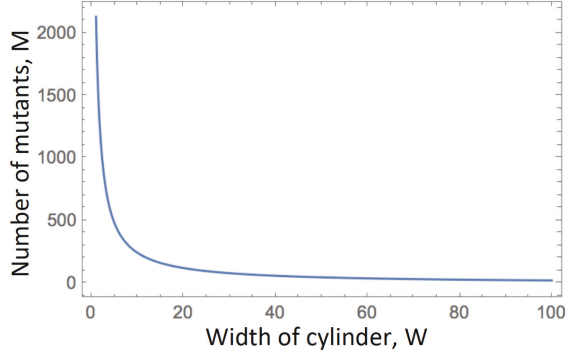


Figure 5: The number of mutants during a 2D flat front expansion decays with the front width. Formula (30) is presented with $N = 10000$ and $u = 5 \times 10^{-5}$.

$$\approx \frac{uWs\mathcal{L}^3}{6} = \frac{us}{6W^2}N^3,$$

where for the approximation, we assumed that $\mathcal{L}s \gg 1$. Also, for this simple calculation to be valid, we need to assume that the wedges created by mutants do not come close to the cylinder's width, W , that is, $\mathcal{L}s \ll W$. In particular, formula (31) can be valid for small values of $W > 1$, but only for mutants that are neutral for practical purposes ($s \ll 1$).

Note that when N is fixed, the total number of cell divisions that the system has undergone is also fixed. The number of mutants however is vastly different depending on the spatial configuration. It is the highest for $W = 1$ (one row of cells) and decreases drastically with the width of the cylinder. This is consistent with the notion that spatial restrictions result in a heightened number of mutants, the 1D space ($W = 1$) being the most spatially restrictive system. The reason for this is that in 1D, a mutant, once created, blocks the whole range of expansion and prevents wild type cells from reproducing. The wider the front, the weaker this effect. Further, we note that in the special case where $W = 1$, mutant advantage does not play a role, and the number of advantageous, neutral, and even disadvantageous mutants is given by the same formula, equation (30).

In the derivations above, a zero death rate of cells was assumed. This means that the colony spreads as a solid mass, where all of the spots in the core are occupied. Including a nonzero death rate does not change the geometric argument presented here, because the only difference now is that the expanding population is "porous", such that the same number of cells occupy a larger number of spots on the grid. Therefore, adding a nonzero death rate does not alter the scaling laws derived here and in the other cases. For this reason, we will present calculations assuming a zero death rate. Numerical calculations in

section 5.2 confirm that scaling laws remain the same in the presence of nonzero death rates.

2D: circular range expansion. Next we turn to the dynamics of neutral mutants on a circle. Let us suppose that the radius of the circle is R and $N = \pi R^2$. The size of the colony increases via surface growth with $N \propto t^2$ and

$$R \propto t.$$

As the range expansion proceeds, the circular layer of radius r will on average give rise to $2\pi r u$ new mutations. Each mutation will result in a wedge expanding outwards. If the new mutation occurred in layer with radius r , the number of mutating cells in layer r is 1. The number of mutants in the next layer is given by $\frac{r+1}{r}$, because under the assumption of mutant neutrality, the fraction of mutants in each new layer of radius $j > r$ (with surface $2\pi j$) should stay constant and equal to $\frac{1}{2\pi r}$. For layer j , the number of mutants is then given by j/r . This gives rise to the following calculation:

$$M_{2D \text{ range}}^{neut} = \sum_{r=1}^R 2\pi r u \sum_{j=r+1}^R \frac{j}{r} = \frac{2}{3} \pi R (R^2 - 1) u \approx \frac{2\pi R^3 u}{3} = \frac{2u}{3\pi^{1/2}} N^{3/2}$$

(the approximation is valid for $R \gg 1$).

For advantageous mutants in a growing 2D circle, the fraction of mutants will grow with each layer:

$$\begin{aligned} M_{2D \text{ range}}^{adv} &= \sum_{r=1}^R 2\pi r u \sum_{j=r+1}^R (1 + (j - (r + 1))s) \frac{j}{r} = \pi R (R^2 - 1) u \left(\frac{2}{3} + \frac{1}{4}s(R - 2) \right) \\ &\approx \frac{\pi R^4 s u}{4} = \frac{s u}{4\pi} N^2, \end{aligned} \tag{32}$$

where we assumed $Rs \gg 1$. For this approximation to be valid, the mutant wedges should not exceed the circumference of the colony. Strictly speaking, this results in the condition $Rs \ll 2\pi R$, that is, $s \ll 1$. For larger values of s , the events where the mutant covers the whole surface of the colony are no longer negligible.

3D flat front. In a 3D space, let us first consider a solid cylinder of constant radius R_0 , where initially the cells are situated as a layer at the bottom of the cylinder, and proceed to grow by adding layers of size πR_0^2 . Each generation contributes $\pi R_0^2 u$ new mutants, and as the colony grows to length $\mathcal{L}$ (and volume $2\pi R_0^2 \mathcal{L}$), we have in the neutral case:

$$M_{3D \text{ flat}}^{neut} = \sum_{i=1}^{\mathcal{L}} 2\pi R_0^2 u \sum_{j=i+1}^{\mathcal{L}} 1 = \pi R_0^2 u \mathcal{L} (\mathcal{L} - 1) \approx \pi R_0^2 u \mathcal{L}^2 = \frac{u}{\pi R_0^2} N^2,$$

which is similar to the 2D flat front expansion. If the mutants are advantageous, then their number will increase from layer to layer, giving rise to conical wedges. This gives rise to the following calculation:

$$\begin{aligned}
M_{3D \text{ flat}}^{adv} &= \sum_{i=1}^{\mathcal{L}} 2\pi R_0^2 u \sum_{j=i+1}^{\mathcal{L}} (1 + (j - (i + 1))s)^2 \\
&= \frac{\mathcal{L}(\mathcal{L} - 1)\pi R_0^2 u}{6} [(\mathcal{L}^2 - 3\mathcal{L} + 2)s^2 + 4(\mathcal{L} - 2)s + 6] \\
&\approx \frac{\pi R_0^2 s^2 u \mathcal{L}^4}{6} = \frac{s^2 u}{6\pi^3 R_0^6} N^4,
\end{aligned}$$

where $\mathcal{L}s \gg 1$ for the approximation, and the approach is valid as long as the wedge radius is smaller than that of the cylinder, $s\mathcal{L} \ll R_0$.

3D range expansion. Next we consider a 3D expanding sphere. For a sphere of radius R , we have $N = 4/3\pi R^3$ and the surface is given by $4\pi R^2$. The size of the colony increases via 3D surface growth with $N \propto t^3$. Each spherical layer of radius r will on average give rise to $4\pi r^2 u$ new mutations. Each mutation will result in a conical wedge expanding outwards. If the new mutation occurred in layer with radius r , the number of mutating cells in layer r is 1. The number of mutants in a layer of radius $j > r$ is given by $(j/r)^2$, because under the assumption of mutant neutrality, the fraction of mutants in each new layer should stay constant (and equal to $\frac{1}{4\pi r^2}$). Therefore, we write:

$$M_{3D \text{ range}}^{neut} = \sum_{r=1}^R 4\pi r^2 u \sum_{j=r+1}^R \left(\frac{j}{r}\right)^2 = \pi R(R^2 - 1)(R + 2/3)u \approx \pi R^4 u = \frac{3^{4/3} u}{\pi^{1/3} 4^{4/3}} N^{4/3}$$

(the approximation is again valid for $R \gg 1$).

If the mutant in a growing 3D sphere is advantageous, the fraction in each layer will increase according to the fitness advantage s and stretch from layer to layer in the same way as for the neutral mutants. We therefore have,

$$\begin{aligned}
M_{3D \text{ range}}^{adv} &= \sum_{r=1}^R 4\pi r^2 u \sum_{j=r+1}^R (1 + (j - (r + 1))s)^2 \left(\frac{j}{r}\right)^2 \\
&= \frac{\pi u R(R^2 - 1)}{90} [(20R^3 - 48R^2 - 5R + 42)s^2 + (72R^2 - 90R - 108)s + 90R + 60] \\
&\approx \frac{2}{9}\pi s^2 u R^6 = \frac{s^2 u}{8\pi} N^2.
\end{aligned}$$

As before, the approximation holds if $Rs \gg 1$. The method assumes that the mutant colony's size in each layer does not come close to the surface area, which amounts to the inequality $s \ll 1$.

Curve	Description	L_w	L_m	D_w	D_m
A	Neutral, no death	0.7	0.7	0	0
B	Neutral, with death	0.7	0.7	0.2	0.2
C	Adv, no death	0.7	0.9	0	0
D	Adv, no death, larger advantage	0.7	1.0	0	0
E	Adv by division, with death	0.7	0.8	0.2	0.2
F	Adv by death	0.7	0.7	0.2	0.1
G	Adv by death, wider front	0.7	0.7	0.2	0.1

Table 3: Simulation parameters for figure 6.

Exponential (non-spatial, mass-action) growth. Finally, for exponentially growing populations, similar formulas could be derived. In particular, for neutral mutants, we have

$$M_{\text{exp}}^{\text{neut}} = Nu \ln N,$$

and for advantageous mutants with advantage α (which is the ratio of the net growth rate of mutants and the net growth rate of wild type cells), we have

$$M_{\text{exp}}^{\text{adv}} = \frac{\alpha}{(\alpha - 1)2^{\frac{\alpha-1}{\alpha}}} N^{\frac{2\alpha-1}{\alpha}},$$

see [25], equation (14c), see also equation (13) for a more general formula.

5.2 Comparison with numerical simulations

We have run numerical simulations to check the results derived here. Figures 6 and 7, which are described in detail below, illustrate the accuracy and applicability of the approximations derived above. They contain colored lines, which are results of numerical simulations, and black “guides for the eye”, which represent power-law functions characterized by the powers predicted by the theory. The figures show that these scaling laws hold over large intervals of N , the population size.

Figure 6 shows results for the case of 2D systems, with flat front expansion presented in panel (a) and range expansion in panel (b). Plotted are the average numbers of mutants (the vertical axis) measured at different population sizes, N (the horizontal axis). Different curves correspond to different parameter values, summarized in table 3.

To simulate flat front expansion (panel (a)), ABM simulations were performed in cylindrical geometry, with an $n \times W$ rectangular domain of width W . We started with an array of W wild type cells at the left boundary of the domain, and imposed periodic boundary conditions in the transversal direction. In each simulation, the cell population was allowed to grow to a size N , and the number of mutant cells at this size was recorded. Such simulations were performed repeatedly, and the average number of mutants was calculated. Simulation runs, in which the total cell population went extinct due to stochastic effects, were ignored.

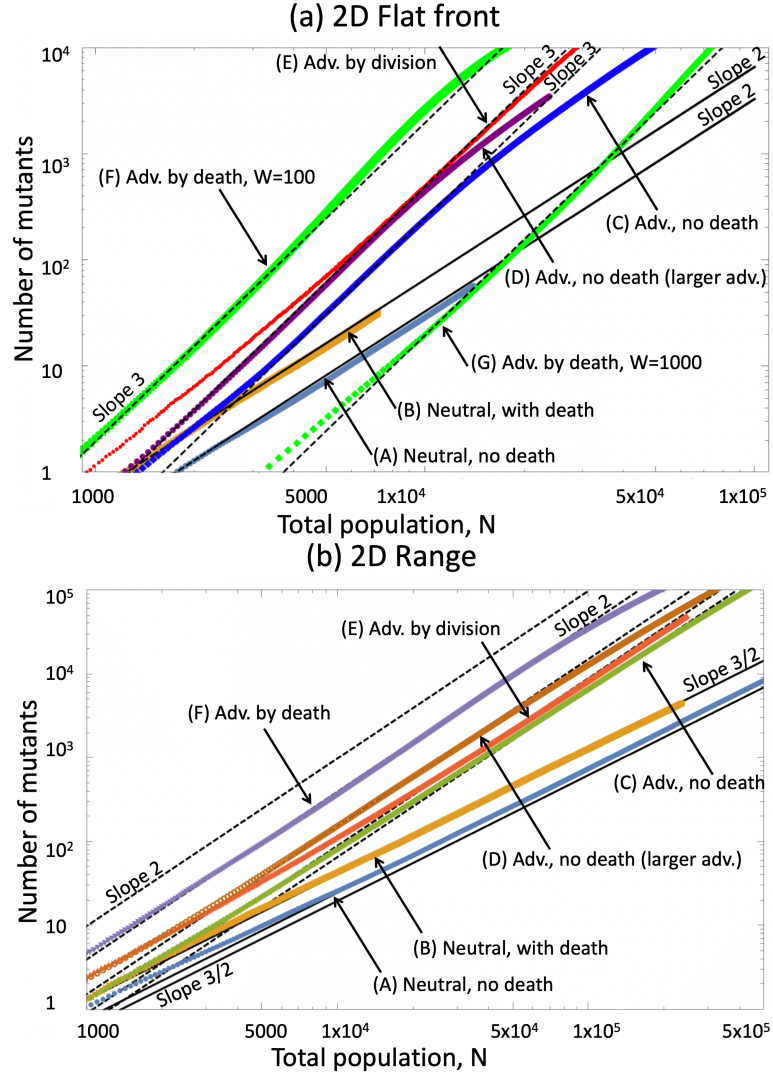


Figure 6: Neutral and advantageous mutants in a 2D colony under (a) a flat front expansion, (b) a range expansion. The number of mutants is plotted as a function of the total population, averaged over 1000 stochastic runs (standard error is too small to see). Cases (A,B) are neutral and corresponding solid black lines are guides to the eye, with (a) slope 2 and (b) slope 3/2 in the log-log plot. Cases (C-G) are advantageous, and the dashed lines are guides to the eye with (a) slope 3 and (b) slope 2. Curves (A-G) are described in table 3. The rest of the parameters are $u = 5 \times 10^{-5}$, $W = 100$ (except G, where $W = 1000$).

Curves (A) and (B) in figure 6(a) represent neutral mutants in the absence (A) and in the presence (B) of cell death. The black solid lines are guides for the eye with slope 2 in the log-log plot, representing the quadratic scaling law (see

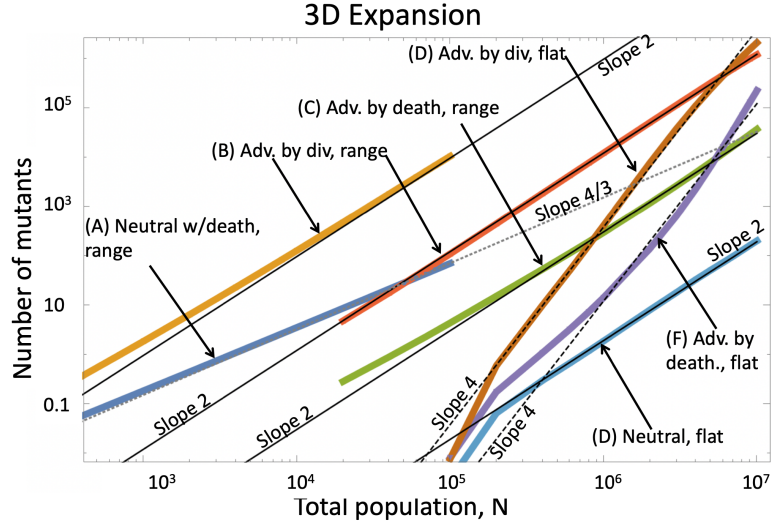


Figure 7: Mutants in the 3D expansion: the average number of mutants is plotted as a function of the total population. (A) Neutral mutants in a range expansion, with the corresponding dotted gray guide to the eye with slope $4/3$ in the log-log plot. (B and C) Advantageous mutants in a range expansion, and the solid lines are guides to the eye with slope 2. (D) Neutral mutants in a 3D flat front expansion, and the solid guide to the eye has slope 2. (E and F) Advantageous mutants in a colony with a 3D flat expansion, the dashed guides to the eye have slope 4. Curve parameters are given in table 4.

table 2). Curves (C-G) represent advantageous mutants and the dashed lines are guides for the eye with slope 3 in the log-log plot, representing the cubic scaling law (see table 2). The different cases of advantageous mutants include systems with and without cell death, and cases where mutants are advantageous by divisions (that is, have a larger division rate and the same death rate, compared with wild type cells), and cases where mutants are advantageous by deaths (that is, have a smaller death rate and the same division rate, compared with wild type cells). The cubic scaling law holds at least for part of the N values for all these cases, see [2] for more details.

To simulate range expansion (figure 6(b)), we performed simulation on a square grid, starting with a single cell in the middle, and letting the population expand outwards. Simulations resulting in population extinction were discarded, and all simulations were stopped before the boundary of the grid was reached. In panel (b), curves (A) and (B) again represent neutral mutants without and with cell death (with solid guides to the eye having slope $3/2$). The rest of the curves again explore advantageous mutants under different assumptions, with dashed guides for the eye having slope 2. In all cases, the scaling laws in table 2 are confirmed.

Curve	Description	L_w	L_m	D_w	D_m	u
A	Neutral, range	0.7	0.7	0.1	0.1	2×10^{-5}
B, yellow	Adv by division, range	0.4	0.8	0.1	0.1	2×10^{-5}
B, red	Adv by division, range	0.7	0.7	0.2	0.2	2×10^{-7}
C	Adv by death, range	0.7	0.7	0.2	0.1	2×10^{-7}
D	Neutral, flat	0.8	0.8	0.1	0.1	2×10^{-7}
E	Adv by division, flat	0.4	0.8	0.1	0.1	2×10^{-7}
F	Adv by death, flat	0.7	0.7	0.2	0.1	2×10^{-7}

Table 4: Simulation parameters for figure 7.

Finally, figure 7 demonstrates numerical results for 3D expanding colonies (with parameters listed in table 4). The different cases considered include neutral and advantageous mutants experiencing 3D flat and range expansion, with advantage realized through differences in division and death rates.

6 Conclusions

In this paper we reviewed some recent results on the behavior of populations in lattice models. Both homogeneous populations (that is, populations consisting of a single type of individuals) and evolving populations (wild type and mutants) were considered, and some simple laws derived that approximate population densities in different dimensionalities, geometries of growth, and lattice types. The methodologies developed here can be extended to other cases, for example, formula (1) of mutant density was derived for a 2D von Neumann grid, but the methodology can be generalized e.g. to the Moore lattice and also to 3D systems.

Our results have further practical applications, for example, for understanding the dynamics of drug resistant mutants in solid tumors [26], or the laws of evolution in bacterial biofilms [27].

The laws described in this paper can be viewed as a way to reduce a complex (spatial, stochastic) process to a small number of important observables, such as the mean equilibrium population density or the expected number of mutants. These quantities are shown to obey some simple rules, which relate these observables to the microscopic (kinetic) parameters of the cellular turnover and to system geometry. Having these simple rules can be very useful, if one, for example, fits a stochastic ABM to a set of data: some parameters can be extracted by applying these laws and solving for the unknown quantities.

Aspects of the theory presented here can be tested in biological experiments. Even though lattice models present a certain simplification of reality, the scaling laws of mutant growth (table 2) are quite versatile, as they have been shown (numerically) to hold in models with different neighborhood structures, and over a large variety of assumptions on cells' kinetic parameters. While some of these laws were tested previously (see e.g. [20]), others, such as 3D flat and 3D range

expansion laws, have not yet been confirmed experimentally.

Having simple laws that connect the macroscopic state with microscopic variables can be useful in interpreting experimental results. The laws in table 1 describe the equilibrium density of populations of cells undergoing a turnover. By comparing the observed population density in experiments with the theoretical predictions, one could simply deduce the division-to-death ratio of cells growing in a given system. This can provide a useful quantitative measure of cells' state. In particular, the differences in cellular density can inform one of the changes in kinetic parameters that arise as a consequence of changes in the microenvironment. Examples of relevant experiments include investigations of the effect of various drug treatments on the cellular turnover.

Finally, our techniques can be extended to describing more complex systems. One potential application is deriving spatial laws of the microbial colony dynamics, such as bacteria growing in a microfluidic trap [28, 29, 30]. Deriving rules of global dynamics of microbial communities from local interaction rules is a task that is similar in spirit to the efforts that were described in the present paper.

References

- [1] Ignacio A Rodriguez-Brenes, Dominik Wodarz, and Natalia L Komarova. Beyond the pair approximation: Modeling colonization population dynamics. *Physical Review E*, 101(3):032404, 2020.
- [2] Dominik Wodarz and Natalia L Komarova. Mutant evolution in spatially structured and fragmented expanding populations. *Genetics*, 2020.
- [3] Motoo Kimura. On the probability of fixation of mutant genes in a population. *Genetics*, 47(6):713, 1962.
- [4] Zaheerabbas Patwa and Lindi M Wahl. The fixation probability of beneficial mutations. *Journal of The Royal Society Interface*, 5(28):1279–1289, 2008.
- [5] Thomas B Kepler and Mihaela Oprea. Improved inference of mutation rates: I. An integral representation for the Luria–Delbrück distribution. *Theoretical population biology*, 59(1):41–48, 2001.
- [6] Anup Dewanji, EG Luebeck, and Suresh H Moolgavkar. A generalized Luria–Delbrück model. *Mathematical biosciences*, 197(2):140–152, 2005.
- [7] Natalia L Komarova, Lin Wu, and Pierre Baldi. The fixed-size Luria–Delbrück model with a nonzero death rate. *Mathematical biosciences*, 210(1):253–290, 2007.
- [8] Qi Zheng. Progress of a half century in the study of the Luria–Delbrück distribution. *Mathematical biosciences*, 162(1-2):1–32, 1999.

- [9] David Bryder, Derrick J Rossi, and Irving L Weissman. Hematopoietic stem cells: the paradigmatic tissue-specific stem cell. *The American journal of pathology*, 169(2):338–346, 2006.
- [10] Jordi Van Gestel, Hera Vlamakis, and Roberto Kolter. Division of labor in biofilms: the ecology of cell differentiation. *Microbial Biofilms*, pages 67–97, 2015.
- [11] Jeremy S Webb, Michael Givskov, and Staffan Kjelleberg. Bacterial biofilms: prokaryotic adventures in multicellularity. *Current opinion in microbiology*, 6(6):578–585, 2003.
- [12] Natalia L Komarova. Spatial stochastic models for cancer initiation and progression. *Bulletin of mathematical biology*, 68(7):1573–1599, 2006.
- [13] Richard Durrett and Stephen Moseley. Spatial Moran models I. Stochastic tunneling in the neutral case. *The annals of applied probability: an official journal of the Institute of Mathematical Statistics*, 25(1):104, 2015.
- [14] Richard Durrett, Jasmine Foo, and Kevin Leder. Spatial Moran models, II: Cancer initiation in spatially structured tissue. *Journal of mathematical biology*, 72(5):1369–1400, 2016.
- [15] Matti Gralka and Oskar Hallatschek. Environmental heterogeneity can tip the population genetics of range expansions. *Elife*, 8:e44359, 2019.
- [16] Bartłomiej Waclaw, Ivana Bozic, Meredith E Pittman, Ralph H Hruban, Bert Vogelstein, and Martin A Nowak. A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity. *Nature*, 525(7568):261–264, 2015.
- [17] Matti Gralka, Fabian Stiewe, Fred Farrell, Wolfram Möbius, Bartłomiej Waclaw, and Oskar Hallatschek. Allele surfing promotes microbial adaptation from standing variation. *Ecology letters*, 19(8):889–898, 2016.
- [18] Jakub Otwinowski and Joachim Krug. Clonal interference and Muller’s ratchet in spatial habitats. *Physical biology*, 11(5):056003, 2014.
- [19] Maxim O Lavrentovich, Mary E Wahl, David R Nelson, and Andrew W Murray. Spatially constrained growth enhances conversional meltdown. *Biophysical journal*, 110(12):2800–2808, 2016.
- [20] Diana Fusco, Matti Gralka, Jona Kayser, Alex Anderson, and Oskar Hallatschek. Excess of mutational jackpot events in expanding populations revealed by spatial luria–delbrück experiments. *Nature communications*, 7:12760, 2016.
- [21] Philip Gerlee and Alexander RA Anderson. An evolutionary hybrid cellular automaton model of solid tumour growth. *Journal of theoretical biology*, 246(4):583–603, 2007.

- [22] Dominik Wodarz, Zhiying Sun, John W Lau, and Natalia L Komarova. Nearest-neighbor interactions, habitat fragmentation, and the persistence of host-pathogen systems. *The American Naturalist*, 182(3):E94–E111, 2013.
- [23] Natalia L Komarova, Leili Shahriyari, and Dominik Wodarz. Complex role of space in the crossing of fitness valleys by asexual populations. *Journal of The Royal Society Interface*, 11(95):20140014, 2014.
- [24] Jesse Kreger, Donovan Brown, Natalia L Komarova, Dominik Wodarz, and Justin Pritchard. The role of migration in mutant evolution in fragmented populations. *bioRxiv*, 2021.
- [25] Yoh Iwasa, Martin A Nowak, and Franziska Michor. Evolution of resistance during clonal expansion. *Genetics*, 172(4):2557–2566, 2006.
- [26] Stuart Horswell, Nicholas Matthews, and Charles Swanton. Cancer heterogeneity and “the struggle for existence”: Diagnostic and analytical challenges. *Cancer letters*, 340(2):220–226, 2013.
- [27] Ehud Banin, Diarmaid Hughes, and Oscar P Kuipers. Bacterial pathogens, antibiotics and antibiotic resistance. *FEMS microbiology reviews*, 41(3):450–452, 2017.
- [28] BR Karamched, W Ott, I Timofeyev, RN Alnahhas, MR Bennett, and K Josić. Moran model of spatial alignment in microbial colonies. *Physica D: Nonlinear Phenomena*, 395:1–6, 2019.
- [29] Simon van Vliet, Christoph Hauert, Kyle Fridberg, Martin Ackermann, and Alma Dal Co. Global dynamics of microbial communities emerge from local interaction rules. *PLoS computational biology*, 18(3):e1009877, 2022.
- [30] Anzhelika Koldaeva, Hsieh-Fu Tsai, Amy Q Shen, and Simone Pigolotti. Population genetics in microchannels. *Proceedings of the National Academy of Sciences*, 119(12):e2120821119, 2022.