Environmental spatial and temporal variability and its role in non-favored mutant dynamics

Suzan Farhang-Sardroodi, Amir H. Darooneh, Mohammad Kohandel, Natalia L. Komarova

Supplementary information

Contents

1 Moran models with temporal and spatial variability

The basic Moran process. We start by defining the classic, non-spatial, Moran process with constant division and death rates [1]. Assume that there are N cells that undergo discrete rounds of divisions and deaths. In each round, one division and one death occur (not necessarily in this order, see below), such that the total number of cells, N, remains constant. In the problems studied here, there are two types of cells, to which we will refer as "wild type" (type A) and "mutant" (type B) cells. In the usual Moran process, all wild type cells are characterized with a constant division rate, r_A, and a constant death rate, d_A. Similarly, mutants have a constant division rate, r_B, and a constant death rate, d_B, where in general, $r_A \neq r_B$ and/or $d_A \neq d_B$. There are two kinds of basic updates that can be formulated (and will be used here): the Birth-Death (BD) and the Death-Birth (DB) processes. It has been reported in the literature that the two exhibited significantly different behavior in the context of game theory [2], in graph-structured populations [3, 4], and even in the simplest formulations of the Moran model [5]. Here we describe the two processes.

- **BD.** First, a cell is selected for division with a probability that is proportional to its division rate. For example, if there are m mutants and $N-m$ wild type cells in the system, the probability that a mutant cell will divide is given by

$$P(\text{mut. div.}) = \frac{mr_B}{(N-m)r_A + mr_B},$$

and the probability that a wild type cell will divide is given by $P(\text{w.t. div.}) = 1 - P(\text{mut. div.})$. A division is followed by a death, with probabilities proportional to cells' death rates. We assume that a cell that just divided is excluded from death during the same update. For example, the probability that a mutant will die, following a division of another mutant, is given by

$$P(\text{mut. death}|\text{mut. div.}) = \frac{(m-1)d_B}{(N-m)d_A + (m-1)d_B},$$

and the probability that a wild type cell will die is given by $P(\text{w.t. death}|\text{mut. div}) = 1 - P(\text{mut. death}|\text{mut. div})$. The newly produced offspring of the cell that divided then replaces the dead cell, and the total population remains at N. The progeny cell retains the type of the parent cells. The probability of an increase in the number of mutants is given by

$$P(\text{mut. div.}) \times P(\text{w.t. death}|\text{mut. div}) = \frac{mr_B}{(N-m)r_A + mr_B} \times \frac{(N-m)d_A}{(N-m)d_A + (m-1)d_B},$$

and the probability of a decrease in the number of mutants is

$$P(\text{w.t. div.}) \times P(\text{mut. death}|\text{w.t. div}) = \frac{(N-m)r_A}{(N-m)r_A + mr_B} \times \frac{md_B}{(N-m-1)d_A + md_B}.$$

- **DB.** Now, the order of events in a single update is reversed: first, a cell is chosen for death (proportionally to its death rate), and then a cell is chosen for division (proportionally to its division rate); the cell that has died cannot participate in divisions during the same update. We have, for example,

$$P(\text{mut. death}) = \frac{md_B}{(N-m)d_A + md_B},$$
$$P(\text{mut. div.}|\text{mut. death}) = \frac{(m-1)r_B}{(N-m)r_A + (m-1)r_B}.$$

The probability that the number of mutants will increase after one update is then given by

$$P(\text{w.t. death}) \times P(\text{mut. div.}|\text{w.t. death}) = \frac{(N-m)d_A}{(N-m)d_A + md_B} \times \frac{mr_B}{(N-m-1)r_A + mr_B},$$

and similarly with the probability of mutant decrease.

Interactions on a network. At the next level of complexity, we assume that cell competition happens only among "neighboring" cells. The notion of a neighborhood is defined by means of a network, whose edges identify the neighbors of each cell. Suppose all the locations are numbered by the index $i \in [1, N]$, and for a given configuration of wild type and mutant cells, we denote their division and death rates as r_i and d_i. We have $r_i \in \{r_A, r_B\}$ and $d_i \in \{d_A, d_B\}$. Now, the probabilities of division and death for a given cell are defined by its own rate and the rates of its neighbors. For example, the division probability for a cell at location k in a BD process is given by

$$\frac{r_k}{\sum_{i \in S_k} r_i},$$

where S_k denotes the set of neighbors of node k. For the death event in the BD update, conditional probabilities are formulated, where the death rates of cells in the neighborhood of a given cell are summed up, excluding the cell that divided in the same update. The total probability of a mutant number increase/decrease is obtained by summing up the probabilities of all the individual update possibilities where a mutant division is followed by a wild type death. Two different networks will be used in this study. One is the complete graph, such that all cells are neighbors of each other. In this case, the probability formulas derived for the basic Moran process apply without change. The second network is a circle, where each cell has exactly two neighbors. These two regular networks in some sense represent the two ends of a spectrum, by the number of neighbors that each cell has. The complete graph leads to the so called "mass-action" dynamics, while a circle corresponds to the most stringent, "nearest neighbor" 1D spatial constraints.

Temporal variability. Motivated by the studies of evolution in variable (random) environments, we formulate the Moran model with temporal randomness. Recall that the processes described above are defined by four constants, which are division and death rates of wild type and mutant cells:

r_A, r_B and d_A, d_B. In the process with temporal randomness, for each time-step, the division values r_A and r_B are chosen from a given probability distribution, see figure 1(a) of the main text for a schematic illustration. In this paper we focus on the case where the probability distributions of r_A and r_B are the same. Similarly, the death rate values d_A and d_B are chosen from a single probability distribution. We are interested in mutant fixation probabilities and times, averaged over all realizations of the rate values.

Note that while our model's "environment" is set up to change at every birth/death event, in reality the event time and the physical time are different. This however does not make a difference in the way how individual organisms experience the environment (in our models). Suppose that time intervals between consecutive birth/death events are distributed exponentially (with the mean of one time unit) and the environmental states are drawn from a certain distribution, every (physical) time unit. If we sample such an environment at time points corresponding to birth/death events, one can show that even though the sequence of environmental conditions thus obtained is different, the distribution remains the same.

Spatial variability. Finally, we consider the problem where randomness is spatial, rather than temporal. In this model, fitness values of each cell are defined by (a) its type and (b) its location. Each realization of the evolutionary process is characterized by a fixed set of wild type fitness values, $r_A^1, \ldots, r_A^N$, and a fixed set of mutant fitness values, $r_B^1, \ldots, r_B^N$, where the superscript is referring to a specific location, and the values r_A^i and r_B^i are iid random variables for $1 \leq i \leq N$. These values, once assigned, remain constant throughout the realization, see figure 1(b) of the main text for a schematic illustration of this model and a comparison with the setting with temporal randomness (panel (a)). Similarly, death rate values, d_A^i and d_B^i are assigned randomly for each realization. For a given realization, whenever a wild type cell happens to be at a location k, it will have the division and death rates r_A^k and d_A^k, respectively. A mutant cell at the same location will have the division and death rates r_B^k and d_B^k. The state of the system at a given time point is defined by the set n_i, $1 \leq i \leq N$, where $n_i = 0$ if a wild type cell is at location i, and $n_i = 1$ otherwise.

For the BD process on the complete graph, the probability for a cell at

location k to divide is given by

$$\frac{n_k r_B^k + (1-n_k) r_A^k}{\sum_{j=1}^N (n_j r_B^j + (1-n_j) r_A^j)}.$$

The probability of a subsequent death at location $i \neq k$ is given by

$$\frac{n_i d_B^k + (1-n_i) d_A^k}{\sum_{j \neq k} (n_j d_B^j + (1-n_j) d_A^j)}.$$

The total probability for an increase in the number of mutants can be calculated by adding the probabilities of all the events where first a mutant cell divides and then a wild type cell dies:

$$\sum_{k=1}^N \frac{n_k r_B^k}{\sum_{j=1}^N (n_j r_B^j + (1-n_j) r_A^j)} \sum_{i \neq k} \frac{(1-n_i) d_A^i}{\sum_{j \neq k} (n_j d_B^j + (1-n_j) d_A^j)}.$$

The probability of mutant decrease is calculated similarly. For the DB process, the order of the events is reversed.

To formulate the problem on a circle, we note that the second event can only occur in the immediate vicinity of the cell involved in the first event. For example, the probability of mutant increase for the BD process on a circle is given by

$$\sum_{k=1}^N \frac{n_k r_B^k}{\sum_{j=1}^N (n_j r_B^j + (1-n_j) r_A^j)} \sum_{i \in \{k-1, k+1\}} \frac{(1-n_i) d_i^k}{\sum_{j \in \{k-1, k+1\}} (n_j d_B^j + (1-n_j) d_A^j)},$$

where we assume periodicity in location numbering, identifying location $N+1$ with location 1 and location 0 with location N.

We are interested in the probability of mutant fixation and the conditional mean fixation time, averaged over all realizations of rate values. In what follows, we first study the model with temporal, and then with spatial randomness. For numerical simulations, unless otherwise specified, we use the two-valued probability distribution of parameter values, where the values $1+\sigma$ and $1-\sigma$ occur with equal probabilities.

2 The effect of temporal randomness on fixation time

The time to fixation results are summarized in table 2 of the main text. We observe that if the mutant behaves as if it is selected for (i.e. its probability of fixation increases), its mean conditional time also increases, and if the mutant's fixation probability decreases, so does its mean conditional fixation time. In other words, if mutants are more likely than neutral to fixate, then in the cases when they do get fixated, the process takes longer on average under temporal randomness compared with a non-random case (see also [6] for an insightful discussion of the relationship between fixation probability and timing). This is related to the phenomenon of balancing selection. Once a mutant population rises from low numbers (which increases the probability of fixation), then eventually the resident population becomes a minority and does not "give up" easily, thus increasing the fixation time.

The reason for this increase in the conditional fixation time can be understood if we consider different scenarios for random division/death rate realizations. For concreteness, let us focus on a BD process with random divisions, and consider different realizations of the sequence $\{r_B^{(n)} - r_A^{(n)}\}$, where the superscript refers to the time-step. From the mutants' prospective, there could be "lucky" and "unlucky" sequences (that is, those sequences that are characterized by an abundance of positive and negative entries, respectively). Such sequences are not likely to result in long fixation times: lucky sequences will lead to a relatively fast mutant fixation, and unlucky sequences are most likely to result in mutant extinction. Instead, it is well-balanced sequences that contribute to long mean fixation times.

In figure 1(a) we plot the fixation time, T, vs the realization-mean mutant advantage,

$$\frac{1}{T}\sum_{n=1}^{T}(r_B^{(n)} - r_A^{(n)}),$$

which is the mean fitness advantage of mutants for a given realization (the averaging is performed over the T time steps). Each point corresponds to a particular realization of the process that ended with mutant fixation (figure 1(a)). Similarly, figure 1(b) shows a graph for the realizations that ended with mutant extinction. We observe that positive mean mutant advantage

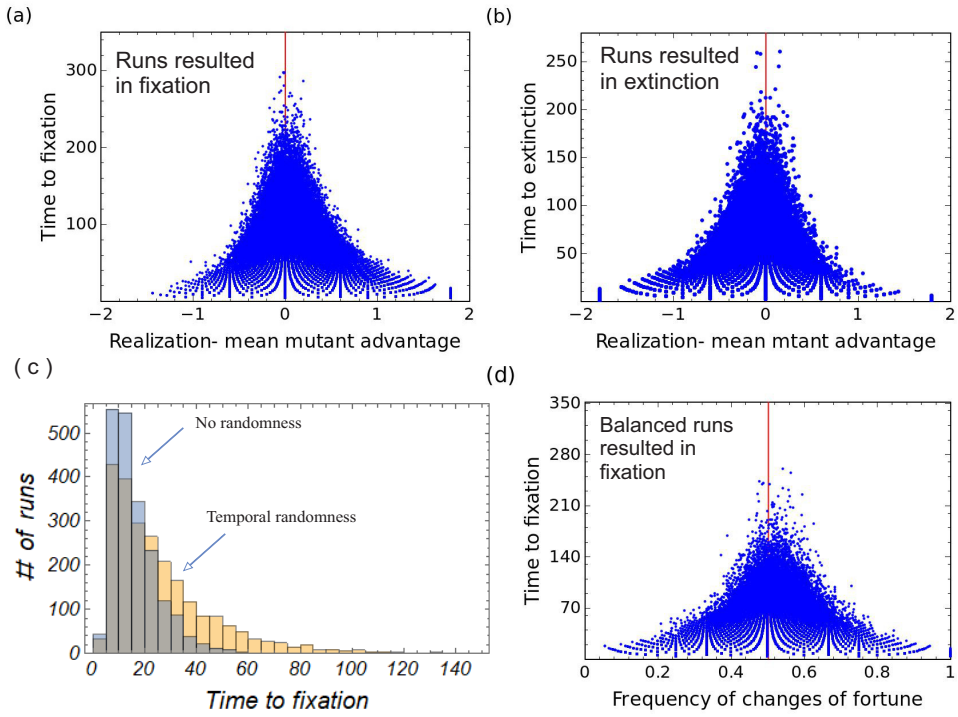

Figure 1: Temporal randomness: time to fixation. (a) Time of fixation vs realization-mean mutant advantage for the runs that resulted in mutant fixation. (b) Time of extinction vs realization-mean mutant advantage for the runs that resulted in mutant extinction. (c) Numerically obtained histograms of fixation times in the presence and in the absence of randomness. (d) Time of fixation vs frequency of changes of fortune, for the balanced runs that resulted in mutant fixation. The total of 10^7 runs were performed with $N = 5$ on a complete graph, and $\sigma = 0.9$ (anti-correlated mutant and wild type division rates, complete graph, BD process with random divisions).

values are stronger associated with mutant fixation and negative ones with extinction. Larger positive and negative mutant advantage values result in shorter fixation times (or extinction). Further, longer trajectories tend to have mean mutant advantage close to zero, which corresponds to what we call "balanced trajectories".

Panel 1(c) shows a numerically obtained probability distribution of fixation times for the system with temporal randomness and compares it to that in the absence of randomness. For the parameters used here the mean conditional fixation time in the non-random case is 16, and that in the presence of randomness is ≈ 25.34. Since it is balanced sequences that largely

contribute to this increase, panel 1(d) considers only runs with the mean mutant advantage in $[-0.1, 0.1]$ (there were about 3×10^5 cases of such balanced runs). We examined the correlation of the time to fixation with the frequency of "changes of fortune" (how many times in the course of a run, mutant advantage changes to disadvantage or back, divided by the sequence length, T). For the particular probability distribution of the division rates, the mean frequency of changes of fortune is $1/2$. We can see that the graph is skewed to the right, revealing that more frequent (than expected) changes of fortune associate with longer fixation times.

3 Time to fixation under spatial randomness

Results for fixation time are summarized in table 4 of the main text. We can see that the relationship between the probability of fixation (table 1) and mean conditional fixation time is different under spatial randomness, compared to the case of temporal randomness, see tables 1 and 2. The probability of fixation increases with any type of randomness on both the complete graph and a circle. For a circle, the time to fixation behaves the same: it increases with randomness. We previously explained this by noticing that randomness can create barriers that are hard to overcome. For the complete graph, randomness in deaths again delays fixation. But randomness in divisions speeds up fixation. This is the only situation where the probability and the time of fixation have different trends as functions of randomness.

In figure 2 we can see that there is a significant difference between the probability distributions of mutant fixation times on the complete graph and on a circle with the same number of spots. Figure 3 shows that different fitness configurations correspond to different fixation times. For both the complete graph (panel (a)) and the circle (panel (b)), we can see that only configurations with 5 or more (out of 9) favorable mutant spots correspond to a significant proportion of fixating runs. This result is intuitive, because configurations with most spots disadvantageous for the mutants will typically result in mutant extinction, and thus do not contribute to the conditional fixation time. For configurations with 5 or more favorable spots, the more favorable spots are allocated to mutants, the shorter the median fixation time. This is not surprising either, because highly advantageous (for mutants) configurations tend to fixate quickly. There is a significant difference however between the complete graph and the circle. In the latter system, the

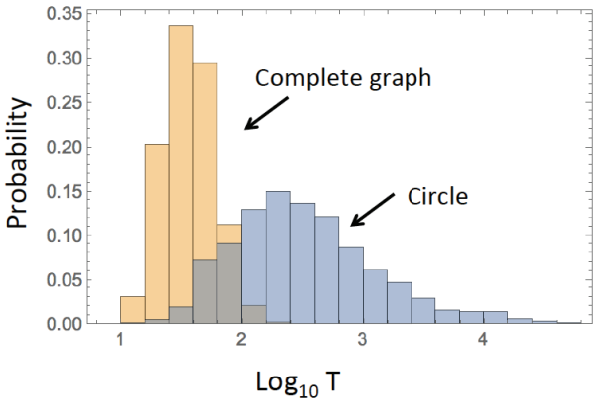

Figure 2: Numerically obtained probability distributions for the conditional time to fixation under spatial randomness. Results for a complete graph and a circle are compared. For these plots we used the BD process with random divisions, the two-valued division rate distribution with $\sigma = 0.9$, $N = 9$, and mutant and wild type division rates were anti-correlated. For each of 2^N fitness configurations, 200 and 50 independent realizations were performed for the complete graph and circle, respectively. Please note the logarithmic horizontal scale.

decay in median fixation time is very fast, as the number of favorable mutant spots increases from 5 to 9 (please note the logarithmic vertical axis). We can see that configurations with 5 favorable mutant spots contribute most significantly to the very long fixation times.

This is why in figure 4 we focus specifically on the configurations that have exactly 5 favorable mutant spots. In panel (a), we separated the runs that had the fastest (the lower quartile) and slowest (the upper quartile) fixation times, and studied exactly what fitness configurations characterized each group. It turned out that in the "slow fixation" group, mutant division rates are favorable (that is, equal $1 + \sigma$) at the original mutant position (that was fixed to be at spot 1) and in the neighboring spots to both sides, while spots opposite of position 1 were unfavorable for mutants (division rates $1 - \sigma$). These spots corresponded to the "dead zones" that we described in [7]. The mutants spread relatively quickly through the favorable island surrounding the initial spot, but then it is very difficult to overcome the dead zone, which results in extremely long fixation times (note that the mean division rates at the initial mutant position are typically high; in the opposite scenario, quick

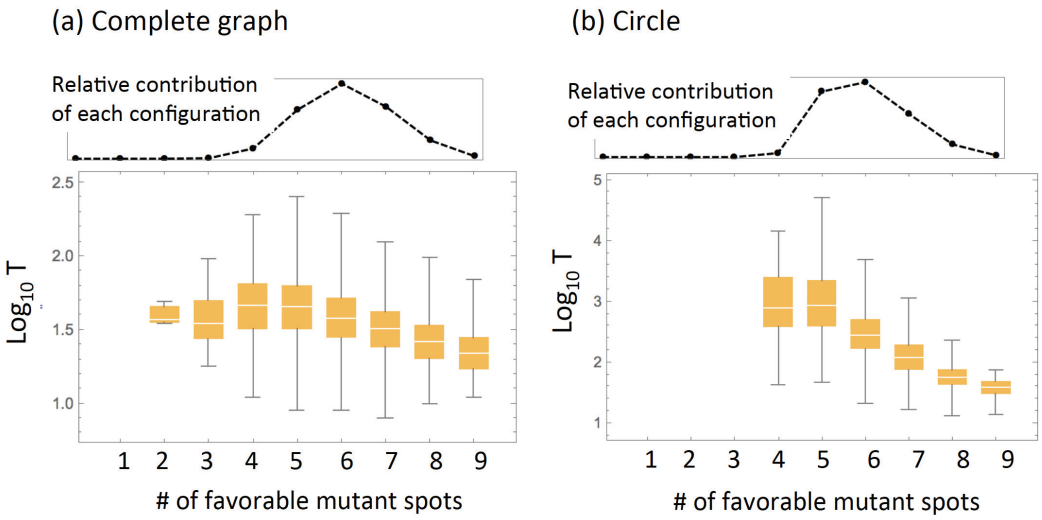

Figure 3: Fixation time for different spatial configurations of the division rates. All configurations were classified by the number of favorable mutant spots, i.e. spots where mutant division rate was $1+\sigma$ (horizontal axis). (a) Complete graph; (b) the circle. Box-and-whisker plots for the fixation times are presented; the graphs on the top indicate relative contribution of each configuration group into the total runs that resulted in mutant fixation. The parameters are the same as in figure 2.

mutant extinction is likely). Now, as we turn to the "fast fixation" group, we can see that while the division rates at the initial mutant location are still quite high (for the same reason), spots away from position 1 are equally likely to be favorable or unfavorable, resulting in the mean division rate of about 1/2. Such configurations correspond to the mutant division rates switching often from favorable to unfavorable, such that there are no large dead zones.

This is exactly what we see in figure 4(b). For each of the configurations containing 5 favorable spots, we calculated the number of switches between $1+\sigma$ and $1-\sigma$, and grouped all runs by the number of such switches. We can see that 2-switch runs (the ones containing large "dead zones") have the largest fixation time, and it decays quickly (please note the logarithmic vertical axis) as the number of switches increases.

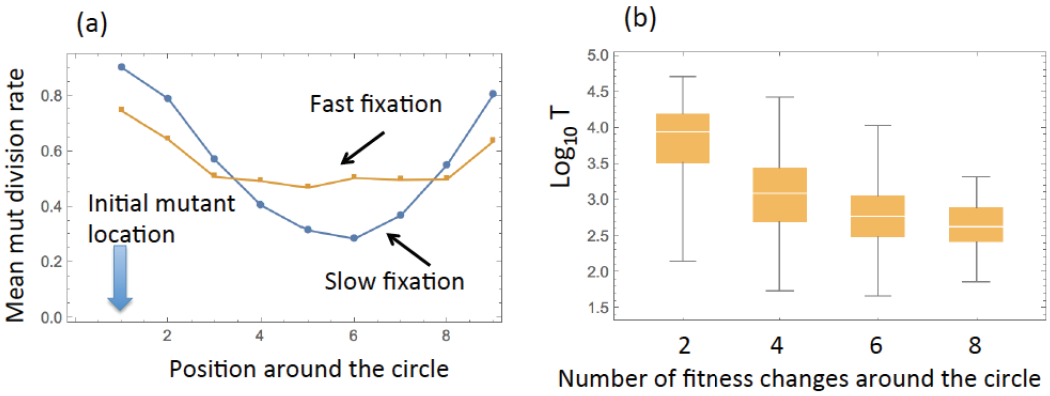

Figure 4: Fixation dynamics on a circle: mutant fixation time for configurations with exactly 5 favorable mutant spots. (a) Using the upper and the lower quartiles of the fixation times, we plotted the mean division rate of mutants for the spots around the circle. "Fast fixation" refers to the runs from the lower quartile and "Slow fixation" to the runs from the upper quartile. (b) Box-and-whiskers plot for the fixation times, split by the number of division rate changes around the circle. The parameters are the same as in figure 2.

References

[1] Patrick Alfred Pierce Moran. Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 54(1):60–71, 1958.

[2] Hisashi Ohtsuki, Christoph Hauert, Erez Lieberman, and Martin A Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441(7092):502, 2006.

[3] Joshua Zukewich, Venu Kurella, Michael Doebeli, and Christoph Hauert. Consolidating birth-death and death-birth processes in structured populations. *PLoS One*, 8(1):e54639, 2013.

[4] Wes Maciejewski. Reproductive value in graph-structured populations. *Journal of Theoretical Biology*, 340:285–293, 2014.

[5] Kamran Kaveh, Natalia L Komarova, and Mohammad Kohandel. The duality of spatial death–birth and birth–death processes and limitations

of the isothermal theorem. *Royal Society Open Science*, 2(4):140465, 2015.

[6] Laura Hindersin and Arne Traulsen. Counterintuitive properties of the fixation time in network-structured populations. *Journal of The Royal Society Interface*, 11(99):20140606, 2014.

[7] Suzan Farhang-Sardroodi, Amirhossein H Darooneh, Moladad Nikbakht, Natalia L Komarova, and Mohammad Kohandel. The effect of spatial randomness on the average fixation time of mutants. *PLoS Computational Biology*, 13(11):e1005864, 2017.