

Do amplifiers of selection maximise average fitness?

Nikhil Sharma* and Arne Traulsen†

*Department of Evolutionary Theory,
Max Planck Institute for Evolutionary Biology,
August-Thienemann-Str. 2, 24306 Plön, Germany.*

(Dated: March 29, 2022)

Abstract

Evolutionary dynamics on graphs has remarkable features: For example, it has been shown that amplifiers of selection exist that – compared to an unstructured population – increase the fixation probability of advantageous mutations, while they decrease the fixation probability of disadvantageous mutations. So far, the theoretical literature has focused on the case of a single mutant entering a graph structured population, asking how the graph affects the probability that a mutant takes over a population and the time until this typically happens. For continuously evolving systems, the more relevant case is when mutants constantly arise in an evolving population. Typically, such mutations occur with a small probability during reproduction events. We thus focus on the low mutation rate limit. The probability distribution for the fitness in this process converges to a steady-state at long times. Intuitively, amplifiers of selection are expected to increase the population’s mean fitness in the steady-state. Similarly, suppressors of selection are expected to decrease the population’s mean fitness in the steady-state. However, we show that another category of graphs, called suppressor of fixation, can attain the highest population mean fitness. The key reason behind this is their ability to efficiently reject deleterious mutants. This illustrates the importance of the deleterious mutant regime for the long-term evolutionary dynamics, something that seems to have been overlooked in the literature so far.

* nsharma@evolbio.mpg.de

† traulsen@evolbio.mpg.de

I. INTRODUCTION

Understanding how spatial structures can affect evolutionary dynamics has been of interest to evolutionary biologists for a long time. More than a decade ago, a framework known as Evolutionary graph theory has been introduced [1]. The primary quantity of interest has been the fixation probability of a mutant on graphs, which is the probability that a mutant with given fitness takes over the rest of the wild-type population [2–6].

Fixation probability is a central concept in evolutionary biology, as it determines the rate of evolution in the low mutation rate regime [7, 8]. Spatial structure tweaks the strength of selection and genetic drift [9, 10]. As a consequence, some graphs have higher probability of fixation than others for a mutant with a given fitness value. Of particular interest are those graphs that increase the fixation probabilities for advantageous mutants and decrease the fixation probabilities for disadvantageous mutants – so called amplifiers of selection. While initially amplifiers seemed to be special structures [1], it turned out that under Birth-death updating, most random networks are amplifiers of selection [11].

However, fixation describes evolutionary dynamics only on a relatively short time scale. In the long run, additional mutants will arise and the population will eventually reach a steady state in terms of fitness. A model that investigates such long evolutionary trajectories in graph structured populations has been missing so far. This problem has been studied for two types [12], but not for the case of continuously arising mutations with a potentially infinite number of types. The neutrality counterpart of this problem has been investigated in [13] where equilibrium properties are shown to be independent of the population structure.

In Evolutionary graph theory, every node of a graph is an individual. If an individual produces an offspring, the offspring is placed in a neighboring node. We focus here on graphs that are undirected and unweighted, such that all neighboring nodes are chosen with the same probability. Graphs can be classified based on their fixation probability as compared to that of the complete graph. The fixation probability of a mutant with fitness f' appearing in a population with fitness f on a complete network C (corresponding to a well-mixed population) is given by

$$\phi_{\mathcal{I}}^C(f', f) = \phi_C(f', f) = \frac{1 - \frac{f}{f'}}{1 - \left(\frac{f}{f'}\right)^N}. \quad (1)$$

In this case, all nodes are equivalent and it does not matter where the mutation occurs. In the general case, the fixation probability depends crucially upon the node where the mutant first arises [14]. In Eq. 1, \mathcal{I} denotes this initialisation scheme, according to which a mutant arises on the network. To be specific, \mathcal{I} is the probability distribution represented by a row vector of size N where element i is the probability with which a mutant arises at node i . In the simplest case, this is a uniform probability – this case is typically referred to as uniform initialisation.

Based on fixation probabilities and the initialisation scheme, graphs typically fall into three categories [11]: Amplifiers of selection, suppressors of selection, and isothermal graphs:

- An *Amplifier of selection* is a graph G that increases the fixation probability of an advantageous mutant and decreases the fixation probability of a disadvantageous mutant compared to a complete graph, $\phi_G^{\mathcal{I}}(f', f) > \phi_C^{\mathcal{I}}(f', f)$ for $f' > f$ and $\phi_G^{\mathcal{I}}(f', f) < \phi_C^{\mathcal{I}}(f', f)$ for $f' < f$.

- A *Suppressor of selection* is a graph G that decreases the fixation probability of an advantageous mutant and increases the fixation probability of a disadvantageous mutant compared to a complete graph, $\phi_G^{\mathcal{I}}(f', f) < \phi_C^{\mathcal{I}}(f', f)$ for $f' > f$ and $\phi_G^{\mathcal{I}}(f', f) > \phi_C^{\mathcal{I}}(f', f)$ for $f' < f$.
- An *Isothermal graph* is a graph G that has the same fixation probability as the complete graph, $\phi_G^{\mathcal{I}}(f', f) = \phi_C^{\mathcal{I}}(f', f)$ for all f' and f . In particular, any graph where the number of links for each node is the same is an isothermal graph for uniform initialisation (in the more general case of weighted graphs, where mutants are placed on neighbouring nodes with different probabilities [15], the temperature of a node i is equal to the sum of the incoming link weights, $\sum_{j=1}^N w_{ji}$).

For later on purposes, we introduce a fourth type [11, 16], Suppressors of fixation:

- A *Suppressor of fixation* is a graph G that reduces the fixation probability for both advantageous and disadvantageous mutants, $\phi_G^{\mathcal{I}}(f', f) < \phi_C^{\mathcal{I}}(f', f)$ for all $f' \neq f$. These graphs are also called reducers of fixation [17].

The definition for isothermal networks has originally been defined for a uniform initialisation scheme. In general, such a definition would depend on the details of the update rule [18]. The classifications for amplifiers and suppressors above have also been developed for uniform mutant initialisation. For other initialisation schemes, for example, temperature initialisation, the classifications become less straightforward, especially near neutrality [19].

II. MODEL

In order to generate any evolutionary dynamics, we must choose an update mechanism. We focus on the Birth-death (Bd) update rule, where first an individual is selected at random, but proportional to fitness. This individual produces an offspring, which is placed in one of the neighbouring nodes. We assume that the offspring is mutated with a small probability μ and identical to its parent with probability $1 - \mu$. The state of a population is represented by a fitness vector $\mathbf{f} = (f_1, f_2, \dots, f_{N-1}, f_N)^T$, where f_i is the fitness of an individual at node i . When an offspring is mutated, we choose its fitness f' from a continuous bounded distribution $\rho(f', f)$, where f is the parent's fitness and $f_{\min} \leq f' \leq f_{\max}$, see Fig. 1 (a).

Instead of looking into the full evolutionary trajectory of the system, we focus only on the fitness distribution in the steady state. As this is difficult for arbitrary mutation rates, we concentrate on the low mutation rate regime here. In this regime, the entire population effectively moves as a point (most of the time) in the fitness space. Typically, every individual in a population has the same fitness. Thus, the state of the population can be labeled by a single fitness value f . This is a good approximation if the time between two successive mutations is sufficiently high that a new mutant either gets extinct or takes over the entire population before the next mutation arises. Thus, our model describes sequential fixation [20]. We are mainly interested in the dynamics of probability density $P_G(f, t)$. One important observation is that mutations do not arise in all nodes with uniform probability, but mostly in those nodes that have more incoming links. This is captured by the idea of “temperature initialisation” [19], where the probability that a mutation arises is proportional to the temperature of a node i , in the case of unweighted networks $\sum_{j=1}^N w_{ji}$, where $w_{ji} = 1$ if a link between j and i exists.

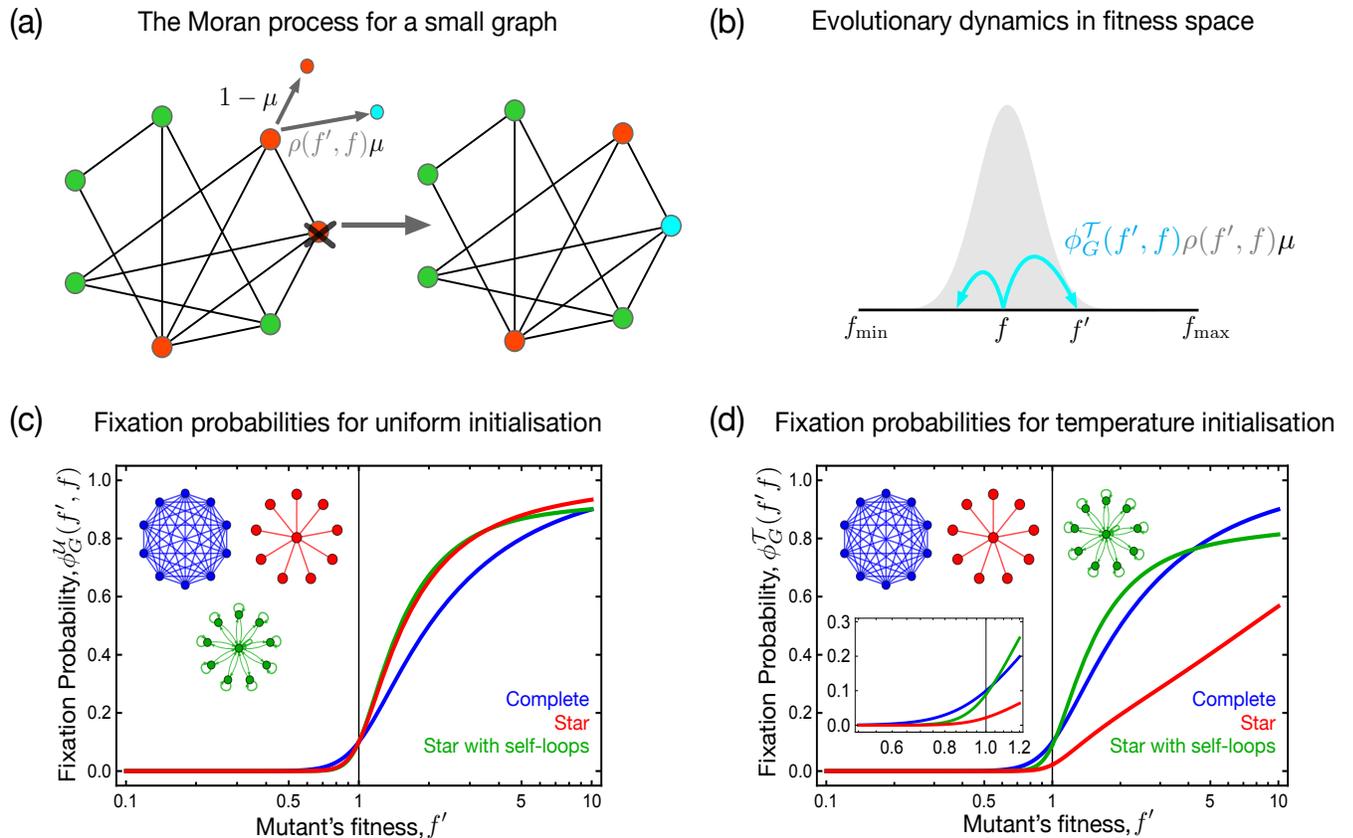


FIG. 1. Moran Birth-death (Bd) with continuous mutation on a graph. (a) The Moran Birth-death (Bd) update mechanism with continuous mutation is shown for a small graph. An individual is selected to reproduce with probability proportional to its fitness. The offspring mutates with probability μ , and its fitness f' is sampled from a distribution $\rho(f', f)$ where f is the parent's fitness. A neighboring individual is then chosen for death at random among the neighboring nodes and the offspring is placed in the empty node. We work in the low mutation rate approximation where μ is very small, such that only a single type is typically present in the population. (b) Evolutionary dynamics in fitness space. For low mutation rates, the evolutionary dynamics effectively becomes a biased random walk on the fitness space. The transition rates depend on the mutation rate μ , the fitness distribution of the offspring $\rho(f', f)$ (shown in grey), and the fixation probabilities $\phi_G^T(f', f)$. (c) The fixation probabilities $\phi_G^U(f', f)$ as function of mutant's fitness f' with $f = 1$ for uniform initialisation for three graphs with $N = 10$: complete, star and (weighted) star with self-loops. For this initialisation scheme, both the star and the star with loops are amplifiers of selection and all fixation probabilities intersect at $f' = f$ and $\phi_G^U(f', f) = N^{-1}$. (d) Fixation probabilities as panel (c), but for temperature initialisation. Now the star with self-loops is only a piecewise amplifier of selection, as it reduces the fixation probability for very large f' . The star is a suppressor of fixation and reduces the fixation probability for all mutant fitness values. Note that even under neutrality, $f' = 1$, the fixation probabilities are different.

With this, we can write down the master equation of the corresponding jump process

[21, 22],

$$\begin{aligned} \frac{\partial P_G(f, t)}{\partial t} &= \int df' \underbrace{\phi_G^{\mathcal{T}}(f, f') \rho(f, f') \mu}_{T_{f \leftarrow f'}} P_G(f', t) \\ &\quad - \int df' \underbrace{\phi_G^{\mathcal{T}}(f', f) \rho(f', f) \mu}_{T_{f' \leftarrow f}} P_G(f, t), \end{aligned} \quad (2)$$

where $T_{f' \leftarrow f}$ is the probability to change the population's fitness from f to f' . $T_{f' \leftarrow f}$ is given by the product of the probability of mutation, μ , the probability that a mutant with fitness f' arises, $\rho(f', f)$, and the probability that such a mutation, arising preferentially in nodes that are replaced often, reaches fixation, $\phi_G^{\mathcal{T}}(f', f)$ (where \mathcal{T} denotes temperature initialisation) The probability $T_{f \leftarrow f'}$ is given by an equivalent argument.

The reason behind using fixation probability under temperature initialisation is that the probability for a mutant to arise on a certain node during reproduction in a homogenous population is proportional to the temperature of that node [19]. The corresponding evolutionary dynamics on the fitness space is a biased random walk with transition rates $T_{f' \leftarrow f}$ and $T_{f \leftarrow f'}$. To derive the steady state, we start from the assumption of detailed balance [22],

$$\int df' T_{f \leftarrow f'} P_G^*(f') = \int df' T_{f' \leftarrow f} P_G^*(f). \quad (3)$$

This condition is fulfilled if we have $T_{f \leftarrow f'} P_G^*(f') = T_{f' \leftarrow f} P_G^*(f)$ for all f, f' . From this, we obtain the steady state [22],

$$P_G^*(f) = \frac{1}{\int df' \frac{T_{f' \leftarrow f}}{T_{f \leftarrow f'}}} = \frac{1}{\int df' \frac{\phi_G^{\mathcal{T}}(f', f)}{\phi_G^{\mathcal{T}}(f, f')} \cdot \frac{\rho(f', f)}{\rho(f, f')}}. \quad (4)$$

If the mutant's fitness distribution $\rho(f', f)$ only depends on the fitness distance between parent and offspring, $\rho(f', f) = \rho(|f' - f|)$, the population's steady-state statistics depends only on the fixation probabilities – in this case, the second factor cancels out. Furthermore, the steady-state probability density for the population's fitness is the same for all isothermal graphs, irrespective of the offspring mutational fitness distribution – the first factor is the same for all isothermal graphs.

To study mutation-selection balance on graphs, we choose the complete graph, the star, and the star with self-loops, i.e. a star where every individual can also be replaced by its own offspring. For finite N , under temperature initialisation a complete graph is an isothermal graph, while the star is a suppressor of fixation and star with self-loops is a piece-wise amplifier of selection [19] (and only for $N \rightarrow \infty$, it becomes an amplifier in a strict sense), see Fig. 1 (c),(d). Our reason for using these graphs is because exact expressions for the corresponding fixation probabilities are known (see App. VB). The fixation probabilities for these three graphs are plotted as a function of mutant's fitness in Fig. 1 (d).

The microscopic Moran-Bd update is costly to simulate for low mutation rates, especially for large graphs. Thus, we use a coarse grained description where we focus on the changes arising from mutations. We use the following Monte-Carlo type algorithm, which measures time steps in terms of mutational events,

- An initial fitness value f between f_{\min} and f_{\max} is assigned to each individual in a population on a graph G .
- At every time step, a mutant f' is drawn from a distribution $\rho(f', f)$.
- The fitness of the entire population is then updated with probability $\phi_G^T(f', f)$ and with probability $1 - \phi_G^T(f', f)$ it remains the same. Note that this step takes into account that mutants tend to arise in different places with different probabilities.
- The last two steps are repeated for a sufficiently long number of time steps until a mutation-selection balance is attained.

We use this algorithm to infer how mutation-selection balance is attained in our three spatial structures. In Fig. 2 (a), the average fitness of the population is plotted as a function of the number of mutations occurred.

III. RESULTS

The star with self-loops, an amplifier of selection, reaches the highest steady state fitness in its mutation-selection balance, see Fig.2 (a). This is expected because, compared to the other graphs, the self-looped star graph is best at fixing beneficial mutants for temperature initialisation. What is surprising is that the star graph, a suppressor of fixation with lower fixation probability for all fitness values of a mutant, not just attains higher steady state fitness than the complete graph (and thus all isothermal graphs), but almost the same balance as the star with loops. We also observe that just like isothermal networks, the star graph with loops takes fewer mutations to reach the balance than the star graphs, which takes many more mutations to reach the steady-state. In the subsequent sections, we will discuss these observations and give an explanation for them.

A. Amplifiers attain higher steady state fitness in mutation-selection equilibrium

Amplifiers of selection are better at fixing beneficial mutants and avoiding the fixation of deleterious mutants than isothermal graphs. They are expected to attain a higher steady state fitness in the mutation-selection balance than isothermal graphs. We verify this expectation in Fig. 2 (b) and (c), which shows the average steady-state fitness of the population for different population sizes. We consider two types of mutational fitness jump distribution, a uniform and a (truncated) Gaussian, centered around the parent's phenotype. We observe that regardless of the mutational fitness distribution, the star graph with loops attains a higher steady state fitness than the complete graph for all considered population sizes. In App. V A, we provide formal a proof that for any mutational fitness distributions, amplifiers of selection have higher steady-state fitness than well-mixed populations and suppressors of selection.

In the limit of very large graphs, $N \rightarrow \infty$, the steady state fitness approaches the maximal possible value, $\langle f^* \rangle \rightarrow f_{\max}$, for both complete as well as self-looped star graph. In fact, f_{\max} becomes an absorbing state for the evolutionary trajectories on these graphs and thus, the fluctuations around the steady state also goes to 0. This happens because for very large N , the fixation probability for the self-looped star under temperature initialisation becomes

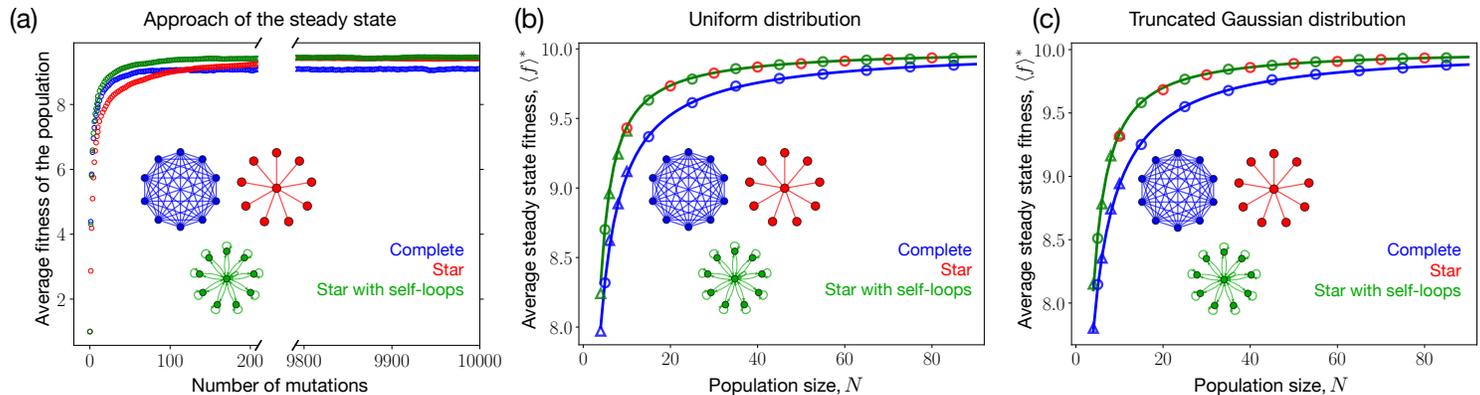


FIG. 2. Average fitness trajectories and steady state fitness for different graphs. (a) The average fitness trajectories for the complete graph, the star graph, and the self-looped star graph with a uniform mutational fitness distribution, starting from a population where every individual has fitness $f = 1$. Although it is a suppressor of fixation, the star graph attains the same steady state fitness in the mutation-selection balance as the self-looped star graph. This is due to its better response to deleterious mutants (see inset of Fig. 1 (d)). (b) and (c) Steady state fitnesses in the mutation-selection balance attained for the self-looped star graph (green) and complete graph (blue) as a function of population size are shown in with (b) uniform and (c) truncated Gaussian mutational fitness jump distribution. The standard deviation for the Gaussian distribution is chosen to be $\sigma = 1$. Solid lines are the numerical solutions of Eq. 4, circles represent simulation points obtained using the Monte-Carlo algorithm, while triangles correspond to microscopic Moran Bd simulations, which are feasible only for very small N . Regardless of the mutational fitness jump distribution, the self-looped star (an amplifier) always attains higher steady state fitness in the mutation-selection balance than isothermal graphs for all (finite) N . Red circles correspond to the steady state fitness attained by the star graph, a suppressor of fixation. From (b) and (c), we find that even for large N , the star graph reaches almost the same steady state fitness as that of the self-looped star graph (parameters $N = 10$, $f_{\min} = 0.1$ and $f_{\max} = 10$, averages over 2000 realisations).

$1 - f^2/f'^2$ for any $f' > f$ and 0 otherwise (see App. VB), whereas from Eq. 1 it follows that for complete graph it becomes, $1 - f/f'$ for $f' > f$ and 0 otherwise. This also leads to the conclusion that the self-looped star needs fewer mutations than the complete graph to reach f_{\max} .

B. Do amplifiers maximise average fitness?

The star graph with no self-loops, a suppressor of fixation, attains almost the same steady state fitness in mutation-selection balance as the star with self-loops, an amplifier of selection, as shown in Fig. 2 (b) and (c). Two main factors drive the steady state fitness: (i) fixing beneficial mutants with high probability and (ii) avoiding the fixation of deleterious mutants. As an amplifier of selection, the star with self-loops is superior in fixing beneficial mutants. But the star without self-loops is much better at avoiding the fixation of deleterious mutants, see inset of Fig. 1(d). In this way, the star without self-loops compensates for its lower fixation probabilities of beneficial mutants. However, this comes at a cost in terms

of the time it takes to reach the steady state. There is additional perspective to this: Let us take two amplifiers A_1 and A_2 , and without any loss of generality, assume that A_1 is a better amplifier than A_2 . That is, A_1 is better in fixing advantageous mutants and avoiding the fixation of deleterious mutants than A_2 . In that case, by the proof given in App. V A, it follows that the average steady state fitness of A_1 is greater than the average steady state fitness of A_2 , i.e. $\langle f \rangle_{A_1}^* > \langle f \rangle_{A_2}^*$. However, if A_2 can be modified, along the lines of [23], such that its fixation probability for deleterious mutants becomes much lower than A_1 , the average steady state fitness $\langle f \rangle_{A_2}^*$ can exceed $\langle f \rangle_{A_1}^*$.

C. Time to reach the steady state

So far, we have studied the steady state average fitness values for different graphs. Now, we discuss the time they take to reach their respective steady states. To estimate these times, we use the concept of mixing times [24]. In a generic stochastic process, the probability distribution $P_G(f)$ defined on a space Ω changes towards its steady state. The mixing time can then be defined by tracking the distance of this evolving distribution to its steady state $P_G^*(f)$ go to zero. Formally, the mixing time t_{mix} is defined as the minimal time when the distance $d(t)$ to the steady state distribution is smaller than a threshold ε ,

$$t_{\text{mix}}(\varepsilon) = \min\{t : d(t) \leq \varepsilon\}, \quad (5)$$

where,

$$d(t) = \frac{1}{2} \sum_{f \in \Omega} |P_G(f, t) - P_G^*(f)|. \quad (6)$$

In Fig. 3 (a), we find that the star graph, a suppressor of fixation, attains the same balance as the self-looped star, but takes many more mutations to reach the steady-state than all other considered graphs. This is natural, as most mutations arising in this graph do not reach fixation. Another interesting observation is that for large N , the self-looped star graph has the smallest mixing time. This contrasts with the typical fixation probability and time relation, where larger fixation probability tends to correlate with higher fixation times [25–28]. Mixing time is difficult to calculate in the general case. But for single rooted graphs with $\phi_G^{\mathcal{I}} \sim \frac{1}{N}$, it can be computed efficiently. In App. V C, the mixing time is computed for this case.

D. Gaussian fitness-phenotype map

Until now, the dynamics were considered in fitness space. However, in many cases mutations occur on the level of an individual’s phenotype. In that case, the state of a population is represented by a phenotype vector $\mathbf{p} = (p_1, p_2, \dots, p_{N-1}, p_N)^T$, where p_i is the phenotypic trait value of the individual at node i . The Moran Bd update still only takes the fitness of individuals into account. Therefore, a phenotype-fitness map $f(p)$ is used to assign a fitness value to a phenotype. The fitness profile of the population then can be denoted by $\mathbf{f}(\mathbf{p}) = (f(p_1), f(p_2), \dots, f(p_N))^T$. Similar to the previous case, when the offspring is mutated, we choose its phenotype p' from a continuous bounded distribution $\rho(p', p)$, where p is the parent’s phenotypic value and $p_{\text{min}} \leq p' \leq p_{\text{max}}$. In this way, the work presented so

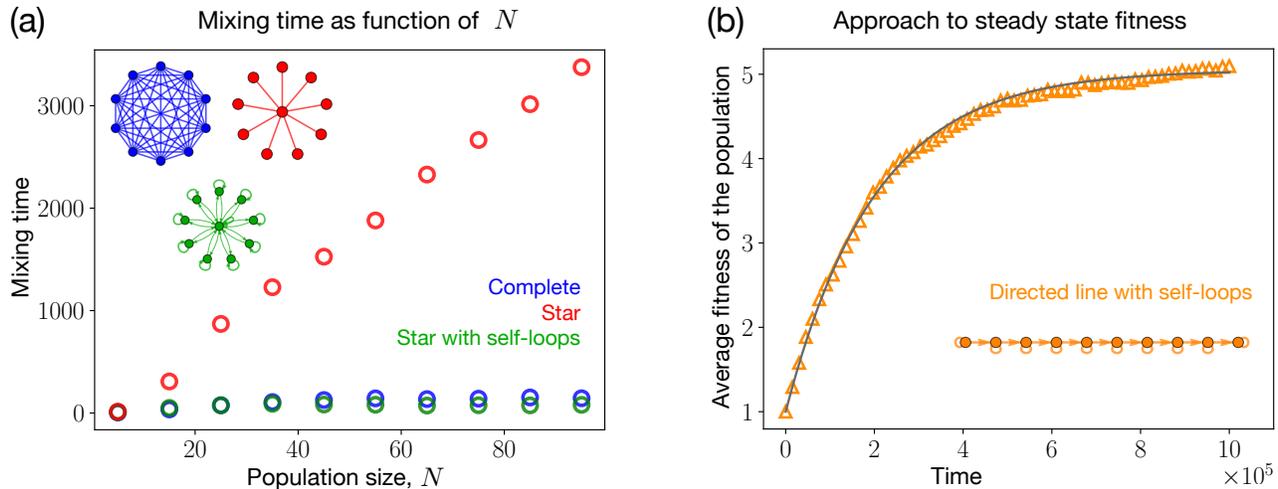


FIG. 3. **Approach to the steady state.** (a) Mixing time for various graphs as a function of population size N measured in number of mutations. Parameters are as in Fig. 2. Mixing times have been obtained using Monte-Carlo algorithm. The star graph takes the longest to reach a steady-state. This is because the star graph, especially for large N , has very small fixation probabilities even for beneficial mutants and thus more attempts to increase fitness are needed to reach the steady state (For the mixing time, we assume $\varepsilon = \frac{1}{4}$ in Eq. 5). (b) Increase of average fitness on the directed line with self-loops. We compare the expression 27 and the corresponding microscopic Moran Bd simulations. Parameters are the same as in Fig. 2, with a mutation probability $\mu = 10^{-4}$. We see a good agreement between simulations (orange symbols) and the analytical result (gray line). From App. VC, we know that the relaxation time for single rooted graphs goes as $\frac{N}{\mu T_r}$. Substituting values for the parameters and the root temperature for the directed line with self-loops, $T_r = \frac{1}{2}$, we get the mixing time to be approximately 2×10^5 . Here, the time is measured in terms of the number of Moran Bd steps. Although the simulation here is performed for the self-looped directed line, the same result is expected for any other single rooted (self-looped) graph as long as μ is sufficiently low. For more details, see App. VC.

far is identical what would obtain by considering a linear phenotype-fitness map, $f(p) = p$. In this section, we study the dynamics of a non-monotonic Gaussian map. In Fig. 4 (a), we show that the order of the steady state fitness in mutation-selection balance for different graphs remains the same as that of the linear phenotype-fitness map. It is interesting to note that the average steady state phenotypic value for all the graphs is the same, see Fig. 4 (b). Nevertheless, we see different steady state average fitness values. Even though all graphs have the same steady state phenotypic value, isothermal graphs are more prone to fluctuations and have a lower steady state average fitness value, see Fig. 4 (c). In a nutshell, this emerges because the average of a function in general is not equal to the function of average, i.e., $\langle f(p) \rangle \neq f(\langle p \rangle)$.

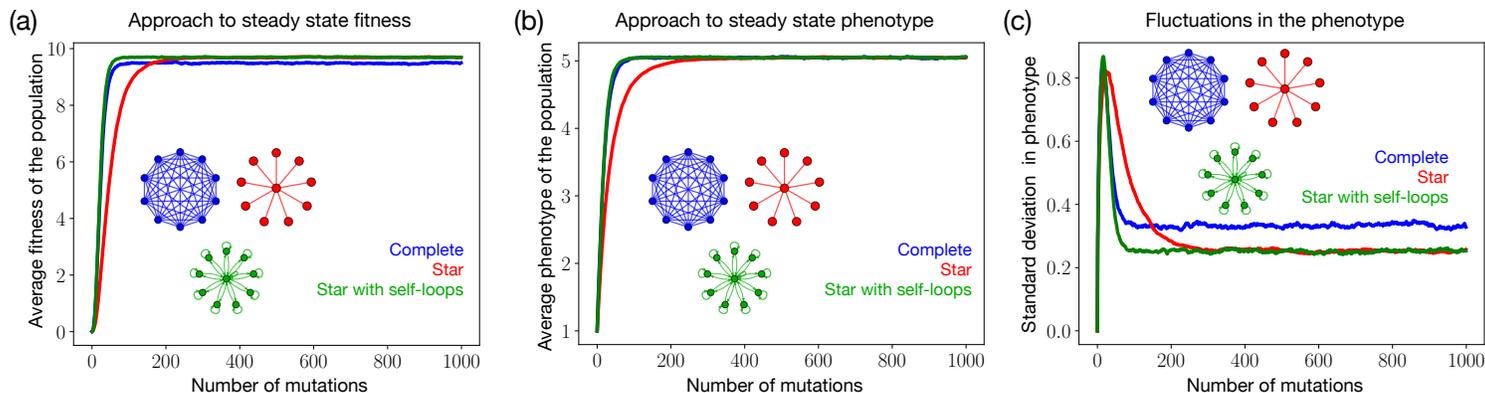


FIG. 4. Gaussian phenotype-fitness map: Fitness and phenotypic trajectories for different graphs. Here, we consider a Gaussian phenotype-fitness map, $f(p) = f_{\text{opt}} \exp(-\frac{(p-p_{\text{opt}})^2}{2\sigma_p^2})$, where p_{opt} is the optimal phenotype where the fitness is maximal and $f(p_{\text{opt}}) = f_{\text{opt}}$. The width of the map is σ_p . (a) As in the case of a linear phenotype-fitness map, the star and the self-looped star have a higher state fitness than isothermal graphs. (b) All graphs converge to an average phenotype corresponding to the fitness maximum, but they have different fitness in steady-state (see panel a). The reason for this is are differences in the fluctuations in steady state. (c) The standard deviation of the phenotypes shows that the complete graph is more prone to fluctuations than the other two graphs (parameters $f_{\text{opt}} = 10$, $p_{\text{opt}} = 5.05$ and $\sigma_p = 1$, all other parameters are as in the previous figures with $\sigma = 0.5$ for the Gaussian mutational jump phenotype distribution).

IV. DISCUSSION

In Evolutionary graph theory, evolutionary dynamics on graph structures have been studied in great detail. While the field has been mostly driven from the mathematical and computational community [1, 2, 18, 27, 29], partly driven by biological systems [16, 30, 31], now there is increasing interest in applying these ideas to experimental systems in microbiology [32, 33].

The prospects of engineering a population structure that can optimise the chances to evolve certain mutations or to observe evolved population structures that minimise the evolution of mutations seem exciting, but these applications call for an extension of the field of Evolutionary graph theory: Most applications implicitly assume that each node is a small population and not all results carry over from graphs of individuals to graphs of sub-populations [34–37]. In addition, the field has focussed so far on fixation probability and fixation time [38–43].

This approach assumes that we can focus on the fate of a single mutant, but it can break down when mutations continuously arise, especially in graphs where the time to fixation or extinction is very high. Moreover, in the case where mutations continuously arise, one has to take into account where they arise. Thus, one needs to work with temperature initialisation, where the definitions of amplifiers and suppressors of selection seem to be less clear-cut.

We developed such a model that takes the continuous supply of mutations into account. We found that the prevention of deleterious mutants from fixing can be more important than increasing the chances of advantageous mutants in order to obtain a higher steady state fitness in the mutation-selection balance. In our case, the star, a suppressor of fixation

for temperature initialisation, beats isothermal graphs and attains almost the same balance as a self-looped star, an amplifier. The cause for this is the ability of the star graph to prevent deleterious mutants much better from fixing than isothermal graphs. The deleterious mutants regime is usually overlooked in the literature while studying fixation probabilities by using large N arguments. However, here we have shown that the deleterious mutants regime is equally important, if not more, even for large N , as the beneficial mutants regime when studying long-term evolutionary dynamics.

Typically, amplifiers of selection have also a higher fixation time [27–29]. Thus, one has to be careful that the assumption of small mutation rate is still fulfilled. The weak mutation approximation works well if the time between two mutation events is much longer than the time to reach fixation or extinction of mutant. Thus, the validity of this approximation depends not only on the mutation rates, but also the fixation times. Previous research has already indicated that graphs with higher fixation times call for lower mutation rates for this approximation to hold [26, 28, 29]. The star graph with self-loops is very restrictive to the low mutation approximation compared to the complete graph, while for the star graph and the complete graph the weak mutation approximation starts to fail at similar mutation rates. This may be unexpected, as for uniform initialisation the fixation time in the star graph is an order of magnitude higher than in the complete graph. But here we observed that for temperature initialisation, the fixation time for the star graph becomes almost the same as that of the complete graph, while the fixation times for self-looped star remain much higher. However, if this assumption of weak mutation rates is fulfilled, structures that allow more mutations to reach fixation will have a smaller mixing time – therefore, amplifiers of selection tend to have a lower mixing time until their steady state fitness is reached.

Due to their ability to reach a high steady state fitness, population structures that suppress selection could be much more interesting in applications than previously thought — suppressing selection may be as relevant as amplifying it. We have shown that in a situation where a dynamic, graph-structured population continuously evolves, the amplification of selection via the promotion of advantageous mutations does not necessarily imply a higher steady state fitness. Instead, in such a process one has to carefully consider also the fate of deleterious mutations, an issue that has so far not been in the focus on Evolutionary graph theory.

V. APPENDIX

A. Amplifiers of selection attain higher steady state fitness in mutation-selection balance than suppressors of selection

Here we prove that amplifiers of selection attain higher steady state fitness in their mutation-selection balance than suppressors of selection. We denote an arbitrary amplifier by A and an arbitrary suppressor by S . From the definitions of amplifiers and suppressors mentioned in the introduction, we have for every $f' < f$ (and arbitrary initialisation scheme \mathcal{I})

$$\phi_A^{\mathcal{I}}(f, f') > \phi_S^{\mathcal{I}}(f, f'), \quad (7)$$

as well as

$$\phi_A^{\mathcal{I}}(f', f) < \phi_S^{\mathcal{I}}(f', f), \quad (8)$$

Combining these two inequalities, we get,

$$\frac{\phi_A^{\mathcal{I}}(f', f)}{\phi_A^{\mathcal{I}}(f, f')} < \frac{\phi_S^{\mathcal{I}}(f', f)}{\phi_S^{\mathcal{I}}(f, f')}, \quad (9)$$

Multiplying the above equation with $\rho(f', f)/\rho(f, f')$ followed by integrating over f' from f_{\min} to f , we obtain,

$$\int_{f_{\min}}^f df' \frac{\phi_A^{\mathcal{I}}(f', f)}{\phi_A^{\mathcal{I}}(f, f')} \cdot \frac{\rho(f', f)}{\rho(f, f')} < \int_{f_{\min}}^f df' \frac{\phi_S^{\mathcal{I}}(f', f)}{\phi_S^{\mathcal{I}}(f, f')} \cdot \frac{\rho(f', f)}{\rho(f, f')}. \quad (10)$$

Similarly, we find,

$$\int_f^{f_{\max}} df' \frac{\phi_A^{\mathcal{I}}(f', f)}{\phi_A^{\mathcal{I}}(f, f')} \cdot \frac{\rho(f', f)}{\rho(f, f')} > \int_f^{f_{\max}} df' \frac{\phi_S^{\mathcal{I}}(f', f)}{\phi_S^{\mathcal{I}}(f, f')} \cdot \frac{\rho(f', f)}{\rho(f, f')}. \quad (11)$$

By making use of the steady-state solution Eq. 4 and Eq. 10, we find an inequality for probability density functions at the boundary f_{\max} of the fitness domain,

$$P_A^*(f_{\max}) > P_S^*(f_{\max}). \quad (12)$$

While Eq. 4 contains temperature initialised fixation probabilities, the same expression follows for the steady state statistics for any arbitrary initialisation. As an example, instead of the mutations taking place during reproduction they could appear spontaneously at any of the nodes. Eq. 4 would then contain uniform initialised fixation probabilities.

Similarly at the fitness boundary f_{\min} , using Eqs. 4 and 11 we have,

$$P_A^*(f_{\min}) < P_S^*(f_{\min}). \quad (13)$$

The very same inequalities, 12 and 13, hold if A (amplifier) or S (suppressor) is replaced by C (complete).

Let us now first prove that amplifiers of selection attain higher steady state fitness in mutation-selection balance than the complete graph. We take the case of an amplifier and the complete graph of the same size. The inequalities 12 and 13 imply that there exists a fitness, \hat{f} , such that,

$$P_A^*(\hat{f}) = P_C^*(\hat{f}). \quad (14)$$

For simplicity here we assumed that there is only one intersection point for the curves $P_A^*(f)$ and $P_C^*(f)$. The proof, however, does not rely on this and can be extended to the general case where more than one intersection point are there. The main idea behind the proof has been sketched in Fig. 5. Let us define the $P_A^*(f)$ as:

$$P_A^*(f) = \begin{cases} P_C^*(f) + \varepsilon_A^<(f), & \text{if } f \leq \hat{f} \\ P_C^*(f) + \varepsilon_A^>(f), & \text{otherwise.} \end{cases} \quad (15)$$

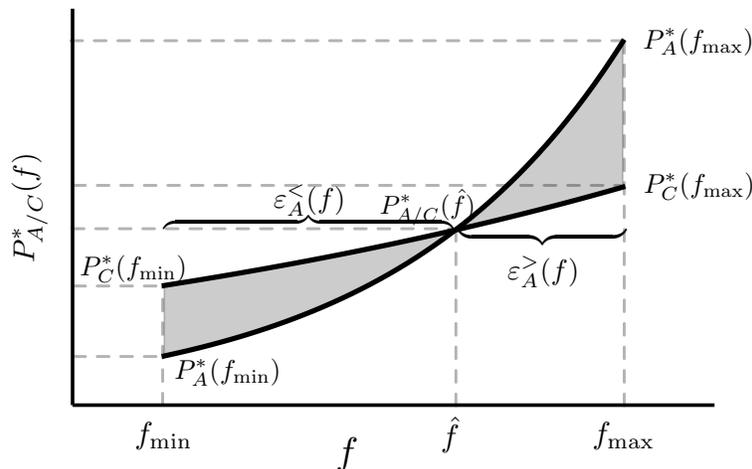


FIG. 5. **Sketch for the proof.** Here we sketch the main idea behind the proof that the average fitness of an amplifier of selection exceeds that of a suppressor of selection or the complete graph in the steady state. Without any loss of generality, we take the case of an amplifier A and the complete graph C . The proof starts by first computing the order of $P_{A/C}^*(f)$ at the fitness boundaries f_{\min} and f_{\max} . It turns out that $P_A^*(f_{\min}) < P_C^*(f_{\min})$ and $P_A^*(f_{\max}) > P_C^*(f_{\max})$. This implies that there exists a fitness point \hat{f} where these probability densities intersect. P_A^* is then decomposed as the sum of P_C^* and the functions $\varepsilon_A^{</>}$, see Eq. 15. By using the normalisation conditions for P_C^* and the properties of functions $\varepsilon_A^{</>}$, Eq. 16, we prove that amplifiers of selection attain higher steady state fitness than the well-mixed population and by extension, the suppressors of selection.

From the normalisation condition of $P_A^*(f)$ and $P_C^*(f)$, it follows that

$$\left| \int_{f_{\min}}^{\hat{f}} df \varepsilon_A^{<}(f) \right| = \int_{\hat{f}}^{f_{\max}} df \varepsilon_A^{>}(f). \quad (16)$$

We are interested in the difference of mean fitness in the steady state,

$$\langle f \rangle_A^* - \langle f \rangle_C^* = \underbrace{\int_{f_{\min}}^{\hat{f}} df f \varepsilon_A^{<}(f)}_{\Gamma_1} + \underbrace{\int_{\hat{f}}^{f_{\max}} df f \varepsilon_A^{>}(f)}_{\Gamma_2}. \quad (17)$$

The first term in the above equation is negative, $\Gamma_1 < 0$, while the second term is positive $\Gamma_2 > 0$. In the following, we show that the magnitude of the term Γ_1 is less than the term Γ_2 . Taking Γ_1 , we find that

$$|\Gamma_1| < \hat{f} \left| \int_{f_{\min}}^{\hat{f}} df \varepsilon_A^{<}(f) \right|. \quad (18)$$

Similarly,

$$\hat{f} \int_{\hat{f}}^{f_{\max}} df \varepsilon_A^>(f) < \Gamma_2. \quad (19)$$

Now, because

$$\hat{f} \left| \int_{f_{\min}}^{\hat{f}} df \varepsilon_A^<(f) \right| = \hat{f} \int_{\hat{f}}^{f_{\max}} df \varepsilon_A^>(f), \quad (20)$$

using inequalities 18 and 19, we find $|\Gamma_1| < \Gamma_2$. Therefore at last, we have

$$\langle f \rangle_A^* > \langle f \rangle_C^*. \quad (21)$$

Following the same procedure, one can show that $\langle f \rangle_C^* > \langle f \rangle_S^*$. This implies that $\langle f \rangle_A^* > \langle f \rangle_S^*$ for all amplifiers of selection A and suppressors of selection S .

B. Fixation probability for the self-looped star graph under temperature initialisation

In this section, we introduce the weighted self-looped star graph that has been used throughout the main text. It is defined by the weighted adjacency matrix

$$\mathbf{w} = \begin{pmatrix} 1 - \delta & \frac{\delta}{n} & \frac{\delta}{n} & \dots & \frac{\delta}{n} \\ \lambda & 1 - \lambda & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \lambda & 0 & \dots & 1 - \lambda & 0 \\ \lambda & 0 & \dots & 0 & 1 - \lambda \end{pmatrix} \quad (22)$$

with $0 < \lambda \leq 1$ and $0 < \delta \leq 1$. Here, w_{ij} is the weight of the link directed from the node i to node j with the center being the node 0. The fixation probability under temperature initialisation for this graph has been derived by using the techniques of martingales [3, 19] and is given by

$$\phi_{G_{n+1}(\lambda, \delta)}^{\mathcal{T}}(f', f) = \frac{1 - \left(\frac{f'}{f}\right)^2}{n\lambda\frac{f'}{f} + \delta} \frac{1}{n+1} \frac{n^3(1-\lambda)\lambda^2 + n^2\lambda\delta\left(\frac{f'}{f} + \lambda\right) + n\lambda\delta\left(\frac{f'}{f} + \delta\right) + (1-\delta)\delta^2}{\left(n\lambda\frac{f'}{f} + \delta\right) \left(\frac{\delta\frac{f'}{f} + n\lambda}{\delta\frac{f'}{f} + n\lambda\left(\frac{f'}{f}\right)^2}\right)^n - \frac{f'}{f} \left(\delta\frac{f'}{f} + n\lambda\right)}, \quad (23)$$

where $G_{n+1}(\lambda, \delta)$ denotes the weighted self-looped star graph with n leaves. In the limit of $n \rightarrow \infty$, when λ and δ are independent of n , Eq. 23 becomes,

$$\lim_{n \rightarrow \infty} \phi_{G_{n+1}(\lambda, \delta)}^{\mathcal{T}}(f', f) = \begin{cases} \left(1 - \left(\frac{f'}{f}\right)^2\right) (1 - \lambda), & \text{if } f' > f, \\ 0 & \text{otherwise.} \end{cases} \quad (24)$$

We use two versions of this weighted graph by first setting $\lambda = \delta = 1$ in the above equation that results in the unweighted star graph which, however, is a suppressor in the limit $n \rightarrow \infty$ under temperature initialisation. This is reflected by $\lim_{n \rightarrow \infty} \phi_{G_{n+1}(\lambda=1, \delta=1)}^{\mathcal{T}}(f', f) = 0$ for all f' and f . Setting instead $\lambda = \frac{1}{n}$ and $\delta = \frac{1}{n^2}$ followed by taking the infinite population size limit of Eq. 23 yields a structure which is an amplifier for $n \rightarrow \infty$. That is,

$$\lim_{n \rightarrow \infty} \phi_{G_{n+1}(\lambda, \delta)}^{\mathcal{T}}(f', f) = \begin{cases} 1 - \left(\frac{f}{f'}\right)^2, & \text{if } f' > f, \\ 0 & \text{otherwise.} \end{cases} \quad (25)$$

C. Mixing time for single rooted graphs

For any one-rooted network with self-loops, we can use Eq. 2 to find an exact expression for the first moment when mutational fitness jump distribution is uniform. To do so, we first note that the fixation probabilities become independent of fitness values, $\phi_G^{\mathcal{T}}(f, f') = \phi_G^{\mathcal{T}}(f', f) = \frac{\mathcal{T}_r}{N}$, where \mathcal{T}_r is the temperature of the root node. We then multiply Eq. 2 with f followed by integrating it over the entire fitness domain. Substituting the fixation probabilities, and using the normalisation condition for $P_G(f', t)$, we obtain a first order ordinary differential equation for $\langle f \rangle$ in time

$$\frac{d\langle f \rangle}{dt} = \frac{\mu \mathcal{T}_r}{N} \left(\frac{f_{\max} + f_{\min}}{2} - \langle f \rangle \right). \quad (26)$$

The average fitness for the case of any single-rooted graph (root with self-loops) with temperature \mathcal{T}_r therefore changes in time as

$$\langle f(t) \rangle = \frac{f_{\max} + f_{\min}}{2} - \left(\frac{f_{\max} + f_{\min}}{2} - \langle f(0) \rangle \right) e^{-\frac{\mu \mathcal{T}_r t}{N}}. \quad (27)$$

From this expression, we find that in the limit $t \rightarrow \infty$, the fitness converges to the average of the fitness domain. Further, the mixing time, the number of mutations required to reach the steady-state, for these graphs scales as $\frac{N}{\mathcal{T}_r \mu}$. Therefore, the mixing time can be modulated by changing the temperature of the root node. We also note that the steady state for single rooted graphs (with self-loops) is independent of the choice of phenotype-fitness map – simply because the fixation probability for these graphs does not depend on fitness.

ACKNOWLEDGMENTS

We are grateful to the members of the Evolutionary Theory department and the Dynamics of Social Behavior group for providing a very good scientific environment facilitating this work. NS sincerely thanks Carsten Fortmann-Grote for his assistance with computer cluster usage.

REFERENCES

- [1] E. Lieberman, C. Hauert, and M. A. Nowak, *Nature* **433**, 312 (2005).
- [2] M. Broom, J. Rychtář, and B. Stadler, *Journal of Interdisciplinary Mathematics* **12**, 129 (2009).
- [3] T. Monk, P. Green, and P. M., *Proceedings of the Royal Society A* **470**, 20130730 (2014).
- [4] L. Hindersin, B. Wu, A. Traulsen, and J. Garcia, *Scientific Reports* **9** (2019).
- [5] Y. P. Kuo, C. Nombela-Arrieta, and O. Carja, *bioRxiv* (2021).
- [6] J. Tkadlec, A. Pavlogiannis, K. Chatterjee, and M. A. Nowak, *PLoS computational biology* **16**, e1007494 (2020).
- [7] M. Kimura, *Nature* **217**, 624 (1968).
- [8] D. M. McCandlish, C. L. Epstein, and J. B. Plotkin, *Theoretical population biology* **99**, 98 (2015).
- [9] M. A. Nowak, *Evolutionary dynamics: Exploring the equations of life* (Harvard University Press, 2006).
- [10] S. Giaimo, J. Arranz, and A. Traulsen, *PLoS Computational Biology* **14**, e1006559 (2018).
- [11] L. Hindersin and A. Traulsen, *PLoS Computational Biology* **11**, e1004437 (2015).
- [12] S. Yagoobi, H. Yousefi, and K. A. Samani, *Physical Review E* **98**, 042301 (2018).
- [13] A. McAvoy, B. Adlam, B. Allen, and M. A. Nowak, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **474**, 20180238 (2018).
- [14] T. Antal, S. Redner, and V. Sood, *Physical Review Letters* **96**, 188104 (2006).
- [15] M. Newman, *Networks* (Oxford university press, 2018).
- [16] L. Hindersin, B. Werner, D. Dingli, and A. Traulsen, *Biology Direct* **11**, 41 (2016).
- [17] B. Allen, C. Sample, P. Steinhagen, J. Shapiro, M. King, T. Hedspeth, and M. Goncalves, *PLoS computational biology* **17**, e1008695 (2021).
- [18] K. Kaveh, N. L. Komarova, and M. Kohandel, *Journal of the Royal Society Open Science* **2** (2015).
- [19] B. Adlam, K. Chatterjee, and M. A. Nowak, *Proceedings of the Royal Society A* **471**, 20150114 (2015).
- [20] D. M. McCandlish and A. Stoltzfus, *The Quarterly review of biology* **89**, 225 (2014).
- [21] C. W. Gardiner, *Handbook of Stochastic Methods*, 3rd ed. (Springer, NY, 2004).
- [22] N. van Kampen, *Stochastic processes in physics and chemistry* (North Holland, 2007).
- [23] A. Pavlogiannis, J. Tkadlec, K. Chatterjee, and M. A. Nowak, *Communications Biology* **1** (2018).
- [24] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov chains and mixing times* (American Mathematical Society, 2009).
- [25] C. Hadjichrysanthou, M. Broom, and J. Rychtář, *Dynamic Games and Applications* **1**, 386 (2011).
- [26] M. Frean, P. Rainey, and A. Traulsen, *Proceedings of the Royal Society B* **280**, 20130211 (2013).
- [27] M. Möller, L. Hindersin, and A. Traulsen, *Communications Biology* **2** (2019).
- [28] J. Tkadlec, A. Pavlogiannis, K. Chatterjee, and M. A. Nowak, *Communications biology* **2**, 1 (2019).

- [29] J. Tkadlec, A. Pavlogiannis, K. Chatterjee, and M. A. Nowak, *Nature Communications* **12**, 4009 (2021).
- [30] J. Cairns, *Nature* **255**, 197 (1975).
- [31] M. A. Nowak, F. Michor, and Y. Iwasa, *Proceedings of the National Academy of Sciences USA* **100**, 14966 (2003).
- [32] P. P. Chakraborty, L. R. Nemzer, and R. Kassen, *bioRxiv* (2021).
- [33] S. van Vliet, C. Hauert, K. Fridberg, M. Ackermann, and A. Dal Co, *PLOS Computational Biology* **18**, e1009877 (2022).
- [34] G. W. Constable and A. J. McKane, *Journal of theoretical biology* **358**, 149 (2014).
- [35] B. Allen, C. Sample, R. Jencks, J. Withers, P. Steinhagen, L. Brizuela, J. Kolodny, D. Parke, G. Lippner, and Y. A. Dementieva, *PLoS computational biology* **16**, e1007529 (2020).
- [36] S. Yagoobi and A. Traulsen, *Scientific Reports* **11**, 1 (2021).
- [37] L. Marrec, I. Lamberti, and A.-F. Bitbol, *Physical Review Letters* **127**, 218102 (2021).
- [38] A. McAvoy and B. Allen, *Journal of Mathematical Biology* **82**, 1 (2021).
- [39] T. Monk, *Journal of Theoretical Biology* **451**, 10 (2018).
- [40] M. Askari and K. A. Samani, *Physical Review E* **92**, 042707 (2015).
- [41] B. Ottino-Löffler, J. G. Scott, and S. H. Strogatz, *Phys. Rev. E* **96**, 012313 (2017).
- [42] D. Hathcock and S. H. Strogatz, *Physical Review E* **100**, 012408 (2019).
- [43] Y. P. Kuo and O. Carja, *bioRxiv* , 2021.06.26.450017 (2021).