

Simulating Ecosystem Evolution via the Generalized Lotka-Volterra Model

Ammaar Saeed

Math 243: Evolutionary Dynamics, Harvard University

May 10, 2022

Abstract

The Moran process is a classical model that describes the stochastic population dynamics of a mutant that arises in isolation. By quantifying the relative fitness of the mutant compared to its source population, it is possible to analytically derive the probability that the mutant drives the source population to extinction, or the fixation probability. Although analytically tractable, real populations are part of a larger ecosystem and interact with a variety of other species in different ways. The generalized Lotka-Volterra (gLV) model provides a method to describe the pairwise interactions between all species of an ecosystem. By combining the stochastic framework of the Moran process with the interaction description of the gLV, it is possible to simulate the impact of interspecies interaction on the fixation probability of a species mutant. Here, we randomly generate ecosystems that give rise to a mutant of a single species and then simulate until the extinction of either the source population or the mutant. We show that, under the conditions tested, the fixation probability can change within a simulated ecosystem of $S = 2$ species, motivating further study of organismic evolution in their proper ecological context.

Introduction

Capturing and quantifying the behavior of evolving systems is a unique challenge in mathematical biology. In the field of evolutionary dynamics, populations consisting of organisms are the smallest evolving units [14]. The evolutionary process proceeds by the random acquisition of novel mutations. This mutant is then subject to the forces of natural selection and random drift. If the mutation is advantageous to the organism, then selection will favor its establishment, while the opposite is true for disadvantageous mutations. However, no deterministic outcome is guaranteed as random factors may eliminate the mutant.

Models for the evolutionary process

Many successful formal approaches within the field frame the evolution of a single organism as natural selection acting solely on the mixed population of the wild-type precursor and a novel mutant. The Moran process is one such approach [13]. The process is a discrete-time Markov chain in which a single mutant individual arises and competes with its source or precursor population. At each time step, one individual is picked for reproduction and one individual is picked for death, with probabilities proportional to relative population size. Importantly, this model can analytically predict the probability that the mutant overtakes the initial source population, or the fixation probability. This quantity is particularly important as it quantifies the rate of evolution [3, 8]. For relative fitness of the mutant r , the fixation probability p is:

$$p = \frac{1 - 1/r}{1 - 1/r^N} \quad (1)$$

A vast amount of work has been done to extend the core ideas of the Moran process to any arbitrary population structure via evolutionary graph theory [9, 16], in which organisms are represented by vertices and their interactions by edges. Although these tools are useful for considering an organism in a vacuum,

natural species interact with a variety of other organisms in complex ecosystems [11]. As such, it is imperative to extend existing models to encapsulate the interactions of real species to test the effects on fixation probability as well as to determine if the same guiding principles continue to apply or must be overhauled with the introduction of complexity.

Interspecies interaction with the Lotka-Volterra model

The Lotka-Volterra model is an ordinary differential equation (ODE) model that is used to describe predator-prey dynamics in an isolated setting (the two species are the only species present) [10, 17]. Taking x to be the population of the prey and y to the population of the predator, the model is typically presented as:

$$\begin{aligned}\frac{dx}{dt} &= \alpha x - \beta xy \\ \frac{dy}{dt} &= \delta xy - \gamma y\end{aligned}\tag{2}$$

The model can be generalized to encapsulate the interactions of a total S species with an interaction matrix \mathbf{A} that describes the impact each species has on the other species and a vector \mathbf{r} of intrinsic growth rates for each species in the absence of other species. A common formulation of the generalized equations is to consider $\mathbf{x}(t)$ the densities of each of the species. The generalized model can be summarized as [1]:

$$\frac{d\mathbf{x}(t)}{dt} = D(\mathbf{x}(t)) (\mathbf{r} + \mathbf{A}\mathbf{x}(t)),\tag{3}$$

where D represents the diagonal of the interaction matrix \mathbf{A} .

Although the gLV model can adequately capture a wide variety of dynamics for many interacting species [6, 15], the model inherently assumes that the interaction matrix is static and thus that populations do not evolve. While this simplifying assumption can be useful for studying populations in the present, mutants can arise and alter the state of an ecosystem with only slightly perturbed interactions compared to its wild-type source. By putting together the two key ideas in these separate models, we propose a model that captures evolution à la the Moran process and accounts for interactions between all species within the environment via the gLV model.

Model and Methods

Stochastic generalized Lotka-Volterra model with evolution

We approach the gLV model under a stochastic framework, modeling ecosystems as a multi-species, fixed population, continuous-time Moran process. This particular process is a subset of continuous-time Markov chains (CTMCs) called a birth-death process for each individual species, wherein an individual from a single species is chosen to replicate proportional to birth rate adjusted by interacting species and another individual from a different species – to maintain this system as a CTMC – is chosen to perish proportional to death rate adjusted by interacting species. The intrinsic birth and death rates of each species are modulated according to an interaction matrix, or an ecological community matrix. Positive effects from one species to another increase a species' birth rate, while negative effects increase a species' death rate.

To consider evolution as in a classical Moran process, a mutant of a single species of interest is introduced into the ecosystem at time $t = 0$. The intrinsic birth and death rate and interspecies interaction parameters are generated by perturbing the relevant parameters of the wild-type as described below. In short, we aim to create a highly correlated but noisy copy of the wild-type species, capturing a realistic mutant. Defining fitness as the difference between birth and death rates, we can compare the relative fitness of the wild-type and mutant species of interest and use the analytical results of the Moran process to derive a theoretical fixation probability for the mutant in the absence of interacting species. We then simulate the gLV model many times as described below until either the wild-type or mutant species of interest fixes and then estimate the fixation probability of the mutant from these trajectories. The code can be found [here](https://github.com/ammsa23/lotka-volterra-evolution) or at this link: <https://github.com/ammsa23/lotka-volterra-evolution>.

Generating birth and death rate parameters

Each species in the model has an intrinsic birth and death rate, which are both drawn independently from a Half-normal distribution with unit variance (Fig. 1a); i.e., the birth rate b and death rate d are sampled as follows:

$$(b, d) \sim (|Z_1|, |Z_2|) \text{ where } Z_1, Z_2 \stackrel{i.i.d.}{\sim} \mathcal{N}(0, 1).$$

To generate the birth and death rates of the mutant, two independent perturbations are drawn from a Uniform distribution on $[-0.1, 0.1]$, yielding highly correlated yet slightly altered birth and death rates for the wild-type and mutant species of interest (Fig. 1a). Based on the initial Half-normal distribution, approximately 95% of draws fall below 2, so we consider the chosen interval to be sizeable and reasonable for perturbation. To account for possible perturbed birth and death rates less than 0, we use the maximum of the perturbed value and 0, so $b_{\text{mutant}} = \max\{0, b_{\text{wild-type}}\}$ and $d_{\text{mutant}} = \max\{0, d_{\text{wild-type}}\}$.

Generating interspecies interaction parameters

We generate the interspecies interaction matrix parameters using two methods with their own set of assumptions. The first method developed by May assumes that all interspecies interactions are independent of each other (Fig. 1b) [12]. This implies that the effect of species A on species B is independent of the effect of species B on species A as well as the effect of species A on species B is independent of the effect of species C on species D. Due to the independence structure, the interaction matrix is not necessarily symmetric. Each of the interaction parameters is drawn from a standard Normal distribution, for which there are $S(S - 1)$ parameters.

The second method developed by Allesina and Tang assumes that there is a dependence between the effect of species A on species B and species B on species A (Fig. 1c) [2]. For example, competing species may both negatively affect each other's relative fitness. Thus, the interactions for a pair of species is sampled from a bivariate Normal distribution mean vector m and covariance matrix Σ as:

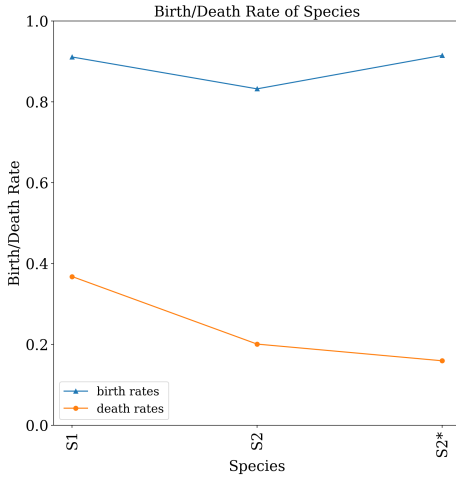
$$m = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ and } \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

The correlation coefficient ρ dictates the interaction dependence. Due to this structure, the interaction matrix must be symmetric. In both cases, the interspecies interactions of the mutant are determined by perturbing the wild-type interaction parameters by independent draws from a Uniform distribution on $[-0.1, 0.1]$.

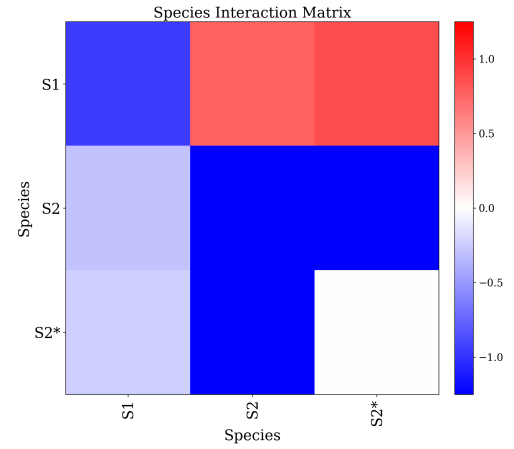
Simulating the stochastic generalized Lotka-Volterra with evolution

We simulate the gLV model described using an algorithm inspired by the Gillespie algorithm [4, 5]. Briefly, after generating random birth and death rate and interspecies interaction parameters for all species and the mutant population, we adjust the birth and death rates by the interspecies interactions. The birth rates are augmented by positive interspecies interactions, while death rates are augmented by negative interspecies interactions. Then, one species gives birth to a single new offspring, with probability proportional to the adjusted birth rate and total number of individuals in that population according to the underlying CTMC. Next, an individual from a species different from the first species dies, with probability proportional to the adjusted death rate and total number of individuals in that population.

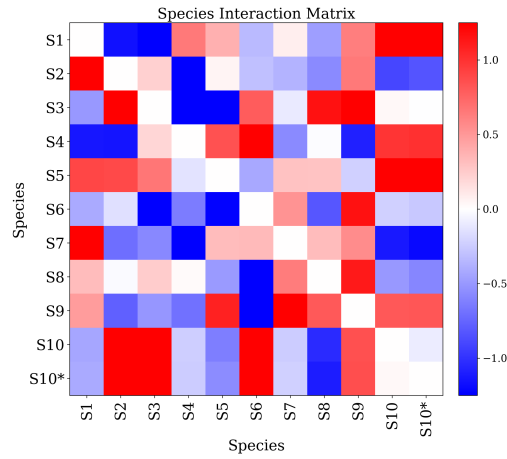
For each of these transitions, the waiting times are drawn from the corresponding Exponential distributions with rate parameters proportional to either birth or death rate and number of individuals part of those populations. This process is repeated until either the wild-type or mutant species of interest fixes, so one of these species must go extinct for a single simulation to reach completion. Random seeds were drawn from 0-32767 to ensure reproducibility of the results.



(a) Birth and Death Rates of Species $S = 2$ (Seed 9411)



(b) Species Interaction Matrix $S = 2$ (Seed 30302)



(c) Species Interaction Matrix $S = 10$ (Seed 27386)

Figure 1: **Example plots for randomly generated parameters for each simulation as described. The generating seeds are included for convenience.** (a) Example plot of the birth and death rates of wild-type, mutant, and peripheral species (b) Example plot of interaction matrix for $S = 2$ species using the May method (c) Example plot of interaction matrix for $S = 10$ species using the Allesina-Tang method

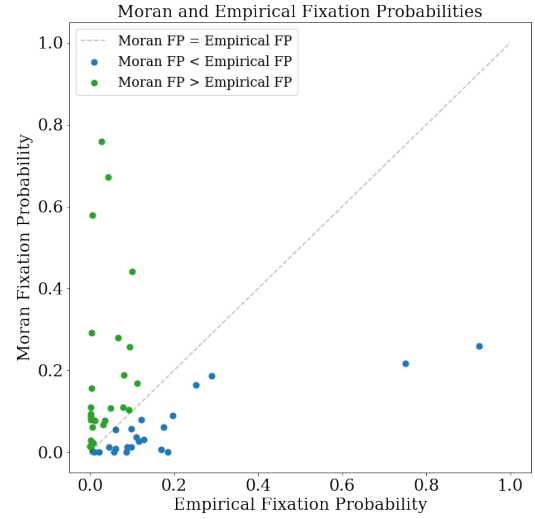
Results

The fixation probability changes with additional species

Under the simulated conditions with both $S = 2$ and $S = 10$ species, we show that the fixation probabilities under the Moran process and the empirically estimated fixation probabilities are different. While this relationship is to be expected given the explicit difference in the modeling choice, we note that there is no general trend in the fixation probabilities as in the fixation probabilities do not globally increase or decrease compared to the Moran process fixation probabilities. Most surprising is that many of these cases occur where the fixation probability is approximately 0 under the Moran process but is non-zero under the current model.

S	N	Method	MFP < EFP	SSDs
2	10	Allesina-Tang	0.5417	0.4682
2	10	May	0.4792	2.5035
2	100	Allesina-Tang	0.1915	1.1338
2	100	May	0.2245	1.8417
10	10	Allesina-Tang	0.7000	1.3714
10	10	May	0.6875	1.0421

Table 1: **Results of randomly generated interaction matrices comparing the analytical Moran fixation probabilities with the empirical fixation probabilities determined by simulation.** The sum of square differences (SSDs) measures the difference between the the two fixation probabilities for each simulation type.



(a) Comparison of Moran and Empirical Fixation Probabilities

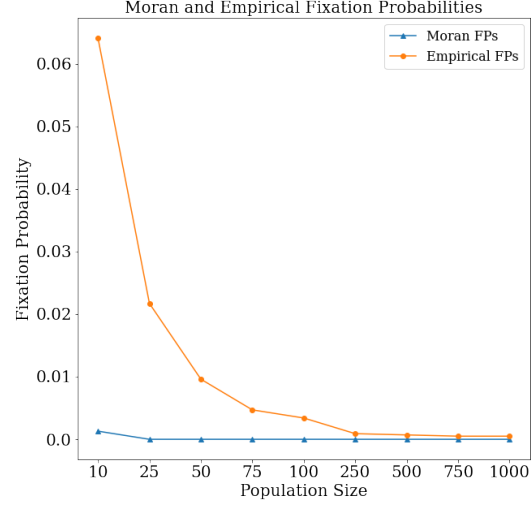
Figure 2: **Plot of Moran and empirical fixation probabilities for $S = 2$ and $N = 10$ using the May method comparing Moran fixation probabilities against empirical fixation probabilities.** Points are colored by comparing the Moran fixation probability against the empirical fixation probability; the dotted grey line indicates where these probabilities are equal.

Interaction matrix generation methods are practically equivalent

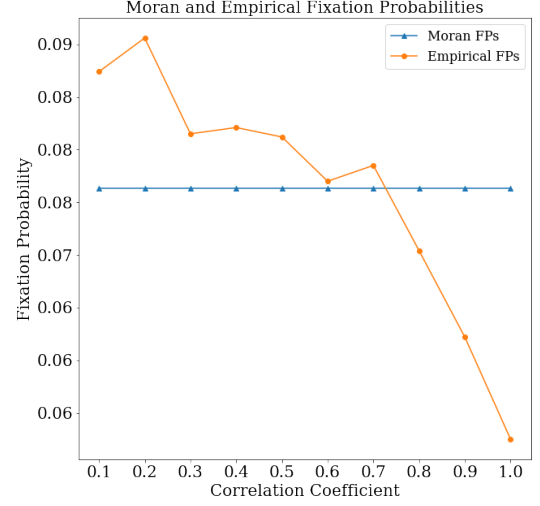
With respect to fixation probability changes, we find for the $S = 2$ and $S = 10$ species cases using both the May and Allesina-Tang methods that the changes in fixation probability are approximately the same. In particular, the proportion of empirical fixation probabilities greater than the Moran process for $S = 2$ and $N = 10$ with the May method is 0.4792 (Fig. 2), while the same using the Allesina-Tang method is 0.5417. Increasing the population size by a factor of 10, we find that the proportion of empirical fixation probabilities greater than that of the Moran process decreases. With $S = 2$ and $N = 100$ under the May method, the proportion is 0.2245, and the same under Allesina-Tang method is 0.1915. When increasing the diversity of the ecosystem to $S = 10$ species and $N = 10$, the proportions increase. With the May method, we find the proportion is 0.6875. The proportion with the Allesina-Tang method is 0.7. These results are summarized in Table 1.

While the overall changes in fixation probability are similar for each using both methods, the true mechanical

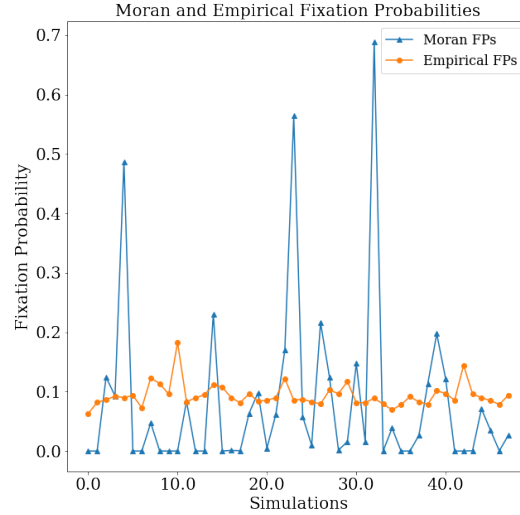
differences for how the fixation probabilities change are quite different. For the $S = 2$ species case, we find that there are generally more drastic changes in the fixation probabilities using the May method. This is particularly true for the $N = 10$ case, for which we compare the sum of square differences (SSDs) and find that this is much larger using the May method (2.5035) compared to the Allesina-Tang method (0.4682). This difference is less pronounced for $S = 2$ and $N = 100$ (1.8417 and 1.1338 for May and Allesina-Tang, respectively) and reverses for the $S = 10$ and $N = 10$ case (1.0421 and 1.3714 for May and Allesina-Tang, respectively) for the trajectories that we have sampled.



(a) Fixation probabilities as a function of population size N (Seed 15944)



(b) Fixation probability as a function of correlation coefficient ρ (Allesina-Tang method; Seed 9855)



(c) Comparison of fixation probabilities for $S = 10$ and $N = 10$ simulations

Figure 3: **Result plots summarizing experimental simulations. Seeds are included for convenience.** (a) Plot of fixation probabilities as population size increases, showing an exponential-like decline (b) Plot of fixation of probabilities as correlation coefficient ρ increases for the Allesina-Tang method (c) Plot of fixation probabilities for many simulations highlighting a near constant empirical fixation probability against fluctuating Moran fixation probabilities

Trends in fixation probability over parameter values

With the current model, there are several testable parameters potentially related to the final fixation probability. Within the Moran process, these parameters are the relative fitness of the mutant r and the total size of the population N . Our model has a couple of tunable parameters that may impact the fixation probability. Using only the May method and thus identical interaction matrices, we test the impact of several parameters in the model, the first of which is the population size for each species N . We find that there is an exponential-like decrease in the fixation probability as the population size increases between $N = 10$ and 1,000 for both the Moran and empirical fixation probabilities (Fig. 3a).

Next, we test the impact of the correlation coefficient ρ in the Allesina-Tang method for creating the generation matrix. As the interactions change, we find that there is an approximately monotonic decrease in the fixation probability as the correlation increases (Fig. 3b). The empirical fixation probability transitions from greater than that of the Moran process between 0.6 and 0.8, indicating that the strength of coupling leads to less favorable outcomes for the mutant individual. Although the decrease is noticeable, we note that the absolute difference between the empirical and Moran fixation probabilities is quite small for all the correlations tested.

Potential limiting behavior of medium-sized ecosystems

From our simulations, the most striking result is the highly consistent fixation probability for the medium-sized ecosystems that we explore. While the Moran fixation probabilities change as we expect based on the variable fitness of the mutant relative to the wild-type, we note that the empirical fixation probabilities do not vary significantly and appear to stabilize at approximately 0.0934 with the May method and 0.0898 with the Allesina-Tang method (Fig. 3c). Both these values are strikingly similar for a variety of ecological interactions between species.

Discussion

Up front, the results of the simulations broadly support our initial hypothesis that other species present in the environment impact the fixation probability. We note that, because there is no global increase or decrease in fixation probability, it is possible to interpret the additional species in the environment as a noise term in the traditional Moran process. Incidentally, this interpretation does fit well with the current implementation of this model as birth and death events can occur in the other species, thus adjusting probabilities of reproduction and death for the wild-type and mutant species of interest.

While we only show singular case studies for each of the parameters of the model, we note that these relationships fall in line with our expectations and so are likely to hold more broadly for most other ecosystem interaction networks. With respect to population size, it follows that the fixation probability would decrease as the population increases as there is a much smaller probability that mutant reproduces (Fig. 3a). The correlation coefficient result also fits in with our expectation as a more highly coupled system of interacting species would strengthen the major interactions within the ecosystem (Fig. 3b). The relationship in the given interaction matrix was highly cooperative between the evolving species of interest and the peripheral species, which was only amplified by the increase in the correlation coefficient, likely making minute differences all the more important.

The highly consistent fixation probability for the medium-sized ecosystems is an interesting result (Fig. 3c). On the one hand, this suggests that the community interactions do not matter for such an ecosystem, except in some relatively rare edge cases. Instead, the fixation probability is dictated by some other parameters of the system that do not involve the individual species. However, it may also be that the majority of interaction matrices of this size are relatively similar, yielding very similar results across the board. It would be interesting to test how this effect scales if at all and to determine which parameters do in fact impact the fixation probability if the effect is truly agnostic to interspecies interactions. It is the opinion of this author that this may strictly be a limitation of the current model and not reflective of evolution in real ecosystems.

While we explore changes in fixation probability, we do not pay particular attention to the interaction matrix,

which details the interactions between members of the ecosystem. For the $S = 2$ species cases, it may be possible to classify the relationships between the species of interest and the interacting species with classical interspecies interaction terms and to find more general patterns across these particular patterns. For the $S = 10$ species cases, this search may be less fruitful as such an interaction contains 90 unique interspecies interactions and would thus be difficult to completely characterize.

In this study, we make several predictions as to how the fixation probability for a mutant changes when introduced into an environment with a finite population and well-defined pairwise interactions between species. It is necessary to test these predictions and others from the current model to validate the model and suggest potential improvements in its assumptions. An appealing experimental mode to verify these results is in diverse bacterial populations. In a test tube, many bacterial species can grow and impact the fitness of surrounding populations in an approximately finite population setting. Such work has been recently published without using evolving populations [7], and a natural extension would be to introduce particular mutants once an equilibrium has been reached. Controlling all other parameters, such as potential neutral mutants within the initial population, makes the experiment technically difficult but all the more intriguing and worthwhile to pursue.

References

- [1] ALLESINA, S. A tour of the generalized lotka-volterra model, May 2020.
- [2] ALLESINA, S., AND TANG, S. Stability criteria for complex ecosystems. *Nature (London)* 483, 7388 (2012), 205–208.
- [3] DESAI, M. M., FISHER, D. S., AND MURRAY, A. W. The speed of evolution and maintenance of variation in asexual populations. *Current biology* 17, 5 (2007), 385–394.
- [4] GILLESPIE, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics* 22, 4 (1976), 403–434.
- [5] GILLESPIE, D. T. Exact stochastic simulation of coupled chemical reactions. *Journal of physical chemistry (1952)* 81, 25 (1977), 2340–2361.
- [6] HIRSCH, M. W. Systems of differential equations which are competitive or cooperative: I. limit sets. *SIAM journal on mathematical analysis* 13, 2 (1982), 167–179.
- [7] HU, J., AMOR, D. R., BARBIER, M., BUNIN, G., AND GORE, J. Emergent phases of ecological diversity and dynamics mapped in microcosms. *bioRxiv* (2021).
- [8] KIMURA, M. Evolutionary rate at the molecular level. *Nature (London)* 217, 5129 (1968), 624–626.
- [9] LIEBERMAN, E., HAUERT, C., AND NOWAK, M. A. Evolutionary dynamics on graphs. *Nature* 433, 7023 (2005), 312–316.
- [10] LOTKA, A. J. Analytical note on certain rhythmic relations in organic systems. *Proceedings of the National Academy of Sciences* 6, 7 (1920), 410–415.
- [11] MAY, R. How many species are there on earth? *Science (American Association for the Advancement of Science)* 241, 4872 (1988), 1441–1449.
- [12] MAY, R. M. Will a large complex system be stable? *Nature (London)* 238, 5364 (1972), 413–414.
- [13] MORAN, P. A. P. Random processes in genetics. *Mathematical proceedings of the Cambridge Philosophical Society* 54, 1 (1958), 60–71.
- [14] NOWAK, M. A. *Evolutionary dynamics : exploring the equations of life*. Belknap Press of Harvard University Press, Cambridge, Mass., 2006.
- [15] SMALE, S. On the differential equations of species in competition. *Journal of mathematical biology* 3, 1 (1976), 5–7.
- [16] TKADLEC, J., PAVLOGIANNIS, A., CHATTERJEE, K., AND NOWAK, M. A. Fast and strong amplifiers of natural selection. *Nature communications* 12, 1 (2021), 4009–4009.
- [17] VOLTERRA, V. *Lecons sur la theorie mathematique de la lutte pour la vie*. Gauthier-Villars, Paris, 1931.