bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Fitness dependence of the fixation-time distribution for evolutionary dynamics on graphs

David Hathcock[1] and Steven H. Strogatz[2]

[1] *Department of Physics, Cornell University, Ithaca, New York 14853, USA*
[2] *Department of Mathematics, Cornell University, Ithaca, New York 14853, USA*

(Dated: December 13, 2018)

Evolutionary graph theory models the effects of natural selection and random drift on structured populations of mutant and non-mutant individuals. Recent studies have shown that fixation times, which determine the rate of evolution, often have right-skewed distributions. Little is known, however, about how these distributions and their skew depend on mutant fitness. Here we calculate the fitness dependence of the fixation-time distribution for the Moran Birth-death process in populations modeled by two extreme networks: the complete graph and the one-dimensional ring lattice, each of which admits an exact solution in the limit of large network size. We find that with non-neutral fitness, the Moran process on the ring has normally distributed fixation times, independent of the relative fitness of mutants and non-mutants. In contrast, on the complete graph, the fixation-time distribution is a weighted convolution of two Gumbel distributions, with a weight depending on the relative fitness. When fitness is neutral, however, the Moran process has a highly skewed fixation-time distribution on both the complete graph and the ring. In this sense, the case of neutral fitness is singular. Even on these simple network structures, the fixation-time distribution exhibits rich fitness dependence, with discontinuities and regions of universality. Applications of our methods to a multi-fitness Moran model, times to partial fixation, and evolution on random networks are discussed.

Significance Statement: Evolutionary graph theory studies the interplay among natural selection, random drift, and the network structure of a population. Fitter mutants tend to reproduce at the expense of wild-type individuals that reproduce less quickly. After a certain amount of time their lineage can take over the entire population. We study a simple stochastic model for the spread of an advantageous mutation and calculate the distribution of times required for the mutants to take over. We find the distribution is often skewed, and establish analytically how the skew depends on the relative fitness of mutants and wild-types, for large networks of individuals connected in a one-dimensional ring or a complete graph. In both cases, the skew jumps discontinuously when fitness is neutral.

Reproducing populations undergo evolutionary dynamics. Mutations can endow individuals with a fitness advantage, allowing them to reproduce more quickly and outcompete non-mutant individuals [1]. Two natural questions arise: If a single mutant individual is introduced into a population, what is the *probability* that the mutant lineage will spread and ultimately take over the population (an outcome known as fixation)? And if fixation occurs, how much *time* does it take?

These questions have been addressed, in part, by evolutionary graph theory, which has advanced our understanding of evolutionary dynamics in structured populations. In particular, fixation probabilities have been studied for various models on various networks [2–12]. The results show that the influence of a network's structure on its fixation probability can be subtle. For instance, under constant selection, a large class of networks known as isothermal graphs have a fixation probability identical to that of a structureless, well-mixed population [2–5], whereas other networks either suppress or amplify selection relative to the well-mixed case [5, 7, 10, 12–14]. When mutations are rare, the fixation probability determines the rate of evolution [1, 15] and can be used to formulate a thermodynamic description of evolution [16, 17].

Other intriguing theoretical studies have focused on constructing optimal structural amplifiers or suppressors to selection [14, 18], or on analyzing more sophisticated dynamics governed by evolutionary games, where selection is frequency dependent rather than constant [6, 11, 19–22]. Parallel efforts have addressed evolutionary dynamics in real-world systems involving cancer growth [23] and response to treatment [24], CRISPR gene drives, [25], and HIV drug resistance [26].

Along with the fixation probability, the fixation time is also of interest, as it too impacts the rate of evolution [15]. Given a model of evolutionary dynamics, one would like to predict the mean, variance, and ideally the full distribution of its fixation times. Of these three quantities, the mean is the best understood. Numerical and analytical results exist for mean fixation times on both deterministic [4, 6, 11, 12, 20, 27–29] and random [28–31] networks. The results reveal curious trade-offs between fixation probability and mean fixation time: network structures that increase the likelihood of fixation may cause the system to take longer to achieve it [12, 15].

Although mean fixation times are important to understand, the information they provide can be misleading, because fixation-time distributions tend to be broad and skewed and hence are not well characterized by their

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

means alone [11, 32–35]. Initial analytical results have determined the asymptotic fixation-time distribution for several simple networks, but only in the analytically convenient limit of infinite mutant fitness [36–38]. For other values of the relative fitness, almost nothing is known. Numerical evidence and preliminary theoretical results suggest that at neutral fitness (when mutants and non-mutants are equally fit), the fixation-time distribution becomes particularly right-skewed [38].

In this paper we investigate the full fitness dependence of fixation times for a simple model of evolutionary dynamics known as the Moran process [39, 40]. In the limit of large network size, we derive asymptotically exact results for the fixation-time distribution and its skew for two network structures at opposite ends of the connectivity spectrum: the complete graph, in which every individual interacts with every other individual; and the one-dimensional ring lattice, in which each individual interacts only with its nearest neighbors on a ring. The resulting fixation-time distributions depend on fitness in rich and contrasting ways.

The specific model we consider is the Moran Birth-death (Bd) process on a network, defined as follows. On each node of the network there is an individual, either mutant or non-mutant. The mutants have a fitness level r designating their relative reproduction rate compared to non-mutants. When $r > 1$, the mutants have a fitness advantage, whereas when $r = 1$ they have neutral fitness. At each time step we choose a node at random, with probability proportional to its fitness, and choose one of its neighbors with uniform probability. The first individual gives birth to an offspring of the same type. That offspring replaces the neighbor, which dies. The model population is updated until either the mutant lineage takes over (in which case fixation occurs) or until the mutant lineage goes extinct (a case we do not consider here).

The quantity of interest is the fixation time, T. Given a single initial mutant on a network of N nodes, and assuming that fixation occurs, how long does it take for the mutants to replace all the non-mutants? As mentioned above, the distribution of fixation times often turns out to be skewed. The skew merges from the stochastic competition between mutants and non-mutants through multiple mechanisms. For instance, when the mutants have neutral fitness, the process resembles an unbiased random walk. Successful runs to fixation will sometimes take long recurrent excursions before the mutants finally sweep through the entire non-mutant population, giving rise to long takeover times and hence a skewed distribution.

Since the neighbor is chosen uniformly, the individuals do not discriminate between mutants and non-mutants during the replacement step of the Moran process. Thus, on certain networks, when there are very few remaining non-mutants, the mutants can waste time replacing each other. This effect produces characteristic slowdowns near the end of the Moran process. These are reminiscent of similar slowdowns seen in a classic problem from probability theory, the coupon collector's problem, which asks: How long does it take to complete a collection of N distinct coupons if one coupon is received randomly (with replacement) at each time step? The problem was first solved by Erdős and Rényi, who proved that for large N, the time to complete the collection has a Gumbel distribution [41]. The intuition for the long slowdowns is clear: When nearly all the coupons have been collected, it can take an exasperatingly long time to collect the final few, because one keeps acquiring coupons that one already has. In fact, an exact mapping onto coupon collection exists for evolutionary processes with infinite fitness [37, 38]. Remarkably, while this correspondence breaks down for finite fitness, the coupon collection heuristic still allows us to write down correct asymptotic fixation-time distributions for non-neutral fitness.

In the following sections we show that for $N \gg 1$, the neutral fitness Moran process on the complete graph and the one-dimensional ring lattice has highly skewed fixation-time distributions, and we solve for their cumulants exactly. For non-neutral fitness the fixation-time distribution is normal on the lattice and a weighted convolution of Gumbel distributions on the complete graph. We begin by developing a general framework for computing fixation-time distributions and cumulants of birth-death Markov chains, and then apply it to the Moran process to prove the results above. We also consider the effects of truncation on the process and examine how long it takes to reach partial, rather than complete, fixation. The fixation-time distributions have rich dependence on fitness level and the degree of truncation, with both discontinuities and regions of universality. To conclude, we discuss extensions of our results to a multi-fitness Moran model and to more complicated network topologies.

GENERAL THEORY OF BIRTH-DEATH PROCESSES

For simplicity, we restrict attention to network topologies on which the probability of adding or removing a mutant in a given time step depends only on the number of existing mutants, not on where the mutants are located on the network. The state of the system can therefore be defined in terms of the number of mutants, $m = 0, 1, \ldots, N$, where N is the total number of nodes on the network. The Moran process is then a birth-death Markov chain with $N + 1$ states, transition probabilities b_m and d_m determined by the network structure, and absorbing boundaries at $m = 0$ and $m = N$. In this section we review and develop general analytical results on the fixation-time distributions for these Markov chains.

On more complicated networks, the probability of adding or removing a mutant depends on the configuration of existing mutants. For some of these networks, however, the transition probabilities can be estimated using a mean-field approximation [31, 37, 38]. Then, to a

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

good approximation, the results below apply to such networks as well.

Eigen-decomposition of the birth-death process

Assuming a continuous-time process, the state of the Markov chain described above evolves according to the master equation. If $\mathbf{p}(t)$ is the probability of occupying each state of the system at time t, then $\mathbf{p}$ evolves according to

$$\dot{\mathbf{p}}(t) = \Omega \cdot \mathbf{p}(t), \tag{1}$$

where Ω is the transition rate matrix, with columns summing to zero. In terms of the transition probabilities b_m and d_m, the entries of Ω can be written as

$$\Omega_{mn} = b_n \delta_{m,n+1} + d_n \delta_{m,n-1} - (b_n + d_n)\delta_{m,n}, \tag{2}$$

where m and n run from 0 to N, $\delta_{m,n}$ is the Kronecker delta, and $b_0 = d_0 = b_N = d_N = 0$. The final condition guarantees the system has absorbing boundaries with stationary states $p_m = \delta_{m,0}$ and $p_m = \delta_{m,N}$ when the population is homogeneous. Thus we can decompose the transition matrix into stationary and transient parts, defining the transient part $\Omega_{\rm tr}$ as in Eq. 2, but with $m, n = 1, \ldots, N-1$. The eigenvalues of $\Omega_{\rm tr}$ are real and strictly negative, since probability flows away from these states toward the absorbing boundaries. To ease notation in the following discussion and later applications, we shall refer to the positive eigenvalues of $-\Omega_{\rm tr}$ as the eigenvalues of the transition matrix, denoted λ_m, where $m = 1, \ldots, N-1$.

From the perspective of Markov chains, the fixation time T is the time required for first passage to state $m = N$, given i initial mutants, $p_m(0) = \delta_{m,i}$. At time t, the probability that state N has been reached (i.e., the cumulative distribution function for the first passage times) is simply $\varphi_i^{-1} p_N(t)$, where φ_i is the fixation probability given i initial mutants. The distribution of first passage times is therefore $\varphi_i^{-1}\dot{p}_N(t) = \varphi_i^{-1} b_{N-1} p_{N-1}(t)$. Since the columns of Ω sum to zero, the right hand side of this equation can be written as $-\varphi_i^{-1}\mathbf{1}\,\Omega_{\rm tr}\,\mathbf{p}(t)$, where $\mathbf{1}$ is the row vector containing all ones.

The solutions to the transient master equation are given by the matrix exponential $\mathbf{p}(t) = \exp(\Omega_{\rm tr} t)\cdot\mathbf{p}(0)$, yielding a fixation-time distribution $-\varphi_i^{-1}\mathbf{1}\Omega_{\rm tr}\exp(\Omega_{\rm tr}t)\cdot \mathbf{p}(0)$. If we assume one initial mutant, $p_m(0) = \delta_{m,1}$, then this can be written as $\varphi_1^{-1} b_{N-1}[\exp(\Omega_{\rm tr}t)]_{N-1,1}$. Since we normalize by the fixation probability, this is exactly the fixation-time distribution conditioned on reaching N. The matrix exponential can be evaluated in terms of the eigenvalues λ_m by taking a Fourier (or Laplace) transform [for details, see Ref. [34]]. For a single initial mutant, the result is that the fixation time T has a distribution $f_T(t)$ given by

$$f_T(t) = \sum_{j=1}^{N-1}\left(\prod_{k=1,k\neq j}^{N-1}\frac{\lambda_k}{\lambda_k - \lambda_j}\right)\lambda_j e^{-\lambda_j t}. \tag{3}$$

Generalizations of this result for arbitrarily many initial mutants have also recently been derived, in terms of eigenvalues of the transition matrix and certain submatrices [34].

The distribution in Eq. 3 is exactly that corresponding to a sum of exponential random variables with rate parameters λ_m. The corresponding cumulants equal $(n-1)!\sum_{m=1}^{N-1}(\lambda_m)^{-n}$. As our primary interest is the asymptotic shape of the distribution, we normalize T to zero mean and unit variance and study $(T-\mu)/\sigma$, where μ and σ denote the mean and standard deviation of T. The standardized distribution is then given by $\sigma f_T(\sigma t + \mu)$. The rescaled fixation time has cumulants

$$\kappa_n(N) = (n-1)!\left(\sum_{m=1}^{N-1}\frac{1}{\lambda_m^n}\right)\Bigg/\left(\sum_{m=1}^{N-1}\frac{1}{\lambda_m^2}\right)^{n/2}, \tag{4}$$

which, for many systems including those considered below, are finite as $N\to\infty$. When the limit exists, we define the asymptotic cumulants by $\kappa_n = \lim_{N\to\infty}\kappa_n(N)$. In particular, because we have standardized our distribution, the third cumulant κ_3 is the skew. Since $\lambda_m > 0$, it is clear from this expression that, for finite N, the skew and all higher order cumulants must be positive, in agreement with analysis on random walks with non-uniform bias [42]. As $N\to\infty$ this is not necessarily true; in some cases the cumulants vanish.

This analytical theory gives the fixation-time distribution and cumulants in terms of the non-zero eigenvalues of the transition matrix. In general the eigenvalues must be found numerically, but in cases where they have a closed form expression the asymptotic form of the cumulants and distribution can often be obtained exactly.

Exact computation of skew via visit statistics

In this section we develop machinery to compute the cumulants of the fixation time analytically without relying on matrix eigenvalues. For this analysis, we specialize to cases where $b_m/d_m = r$ for all m, relevant for the Moran processes considered below. These processes can be thought of as biased random walks overlaid with non-constant waiting times at each state.

It is helpful to consider the Markov chain conditioned on hitting N, with new transition probabilities $\tilde{b}_m$ and $\tilde{d}_m$ so that the fixation probability $\varphi_i = 1$. If X_t is the state of the system at time t, then $\tilde{b}_m = \mathcal{P}(X_t = m \to X_{t+1} = m+1 | X_\infty = N)$ with $\tilde{d}_m$ defined analogously. We derive explicit expressions for $\tilde{b}_m$ and $\tilde{d}_m$ in *SI Abstract*, S1. Conditioning is equivalent to a similarity transformation on the transient part of the transition matrix: $\tilde{\Omega}_{\rm tr} = S\,\Omega_{\rm tr}\,S^{-1}$, where S is diagonal with $S_{mm} = 1 - 1/r^m$. Furthermore, since $b_m/d_m = r$, we can decompose $\Omega_{\rm tr} = \Omega_{\rm RW} D$, where D is a diagonal matrix, $D_{mm} = b_m + d_m$, that encodes the time spent in each state and $\Omega_{\rm RW}$ is the

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

transition matrix for a random walk with uniform bias,

$$[\Omega_{\mathrm{RW}}]_{nm} = \frac{r}{1+r}\delta_{m,n+1} + \frac{1}{1+r}\delta_{m,n-1} - \delta_{m,n}. \quad (5)$$

Applying the results of the previous section, the fixation-time distribution of the conditioned Markov chain is $f_T(t) = -\mathbf{1}\tilde{\Omega}_{\mathrm{tr}}\exp(\tilde{\Omega}_{\mathrm{tr}}t)\mathbf{p}(0)$. The corresponding characteristic function is

$$E[\exp(i\omega T)] = \mathbf{1}\tilde{\Omega}_{\mathrm{tr}}(i\omega + \tilde{\Omega}_{\mathrm{tr}})^{-1}\mathbf{p}(0). \quad (6)$$

Taking derivatives of this expression gives the moments of T,

$$E[T^n] = (-1)^n n!\mathbf{1}\tilde{\Omega}_{\mathrm{tr}}^{-n}\mathbf{p}(0), \quad (7)$$

in terms of $\tilde{\Omega}_{\mathrm{tr}}^{-1} = D^{-1}S\Omega_{\mathrm{RW}}^{-1}S^{-1}$. This inverse has a nice analytical form because S and D are diagonal and Ω_{RW} is tridiagonal Toeplitz. We call this approach *visit statistics* because the elements V_{ij} of $V = -S\Omega_{\mathrm{RW}}^{-1}S^{-1}$ encode the average number of visits to state i starting from state j.

Each power of $\tilde{\Omega}_{\mathrm{tr}}$ in Eq. 7 produces products of $(b_i + d_i)$ that arise in linear combinations determined by the visit numbers V_{ij}. Therefore, the cumulants of the fixation time have the general form

$$\kappa_n(N) = \frac{\displaystyle\sum_{i_1,i_2,\ldots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2\cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})}}{\left(\displaystyle\sum_{i,j=1}^{N-1}\frac{w^2_{ij}(r,N)}{(b_i+d_i)(b_j+d_j)}\right)^{n/2}}, \quad (8)$$

where $w^n_{i_1 i_2\cdots i_n}(r,N)$ are weighting factors based on the visit statistics of the biased random walk. In *SI Abstract*, S2, we give a detailed derivation of Eq. 8 and compute explicit expressions for $w^2_{ij}(r,N)$ and $w^3_{ijk}(r,N)$. This expression for the cumulants is easier to handle asymptotically than the closed form recursive solution for fixation-time moments [33], and can be useful even without explicit expressions for $w^n_{i_1 i_2\cdots i_n}(r,N)$. Estimating these sums allows us to compute the asymptotic fixation time cumulants exactly.

Recurrence relation for distribution moments

Evaluation of the eigenvalues of the transition matrix for large systems can be computationally expensive, with the best algorithms having run times quadratic in matrix size. Numerical evaluation of the cumulants $\kappa_n(N)$ given by Eq. 8 is even worse, as it requires summing $\mathcal{O}(N^n)$ elements. If only a finite number of fixation time cumulants (and not the full distribution) are desired, there are better numerical approaches. Using an approach well known in the probability theory literature [43], we derive a recurrence relation which allows for numerical moment computation with run time linear in system size N. For completeness we provide the full derivation of the reccurence for the fixation-time skew in *SI Abstract*, S3.

ONE-DIMENSIONAL LATTICE

We now specialize to Moran Birth-death (Bd) processes, starting with the one-dimensional (1D) lattice. We assume periodic boundary conditions, so that the N nodes form a ring. The mutants have relative fitness r, meaning they give birth r times faster, on average, than non-mutant individuals do.

Starting from one mutant, suppose that at some later time m of the N nodes are mutants. On the 1D lattice, the population of mutants always forms a connected arc, with two mutants at the endpoints of the arc. Therefore, the probability b_m of increasing the mutant population by one in the next time step is the probability of choosing a mutant node at an endpoint to give birth, namely $2r/(rm + N - m)$, times the probability $1/2$ that the neighboring node to be replaced is not itself a mutant. (The latter probability equals $1/2$ because there are two neighbors to choose for replacement: a mutant neighbor on the interior of the arc and a non-mutant neighbor on the exterior. Only the second of these choices produces an increase in the number of mutants.) By multiplying these probabilities together we obtain

$$b_m = \frac{r}{rm+N-m}, \qquad d_m = \frac{1}{rm+N-m}, \quad (9)$$

where the probability d_m of decreasing the mutant population by one is found by similar reasoning. These quantities play the role of transition probabilities in a Markov transition matrix. The next step is to find the eigenvalues of that matrix.

Neutral fitness

First we work out the eigenvalues for the case of neutral fitness, $r = 1$. In this case, the transition probabilities are equal, $b_m = d_m = 1/N$, and independent of m. Therefore, the Moran process is simply a random walk, with events occurring at a rate of $2/N$ per time step. The associated transition matrix is tridiagonal Toeplitz, which has known eigenvalues given by

$$\lambda_m = \frac{1}{N} - \frac{1}{N}\cos\left(\frac{m\pi}{N}\right), \quad m = 1, 2, \ldots, N-1. \quad (10)$$

Applying Eq. 4 and computing the leading asymptotic form of the sums $S_n = \sum_{m=1}^{N-1}(\lambda_m)^{-n}$ (see *SI Abstract*, S4), we find that as $N \to \infty$, the fixation-time distribution has cumulants

$$\kappa_n = (n-1)!\,\frac{\zeta(2n)}{\zeta(4)^{n/2}}, \quad (11)$$

where ζ denotes the Riemann zeta function. In particular, the skew $\kappa_3 = 4\sqrt{10}/7 \approx 1.807$, as calculated by Ottino-Löffler et al. [38] via martingale methods. The largeness of the skew stems from the recurrent property

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

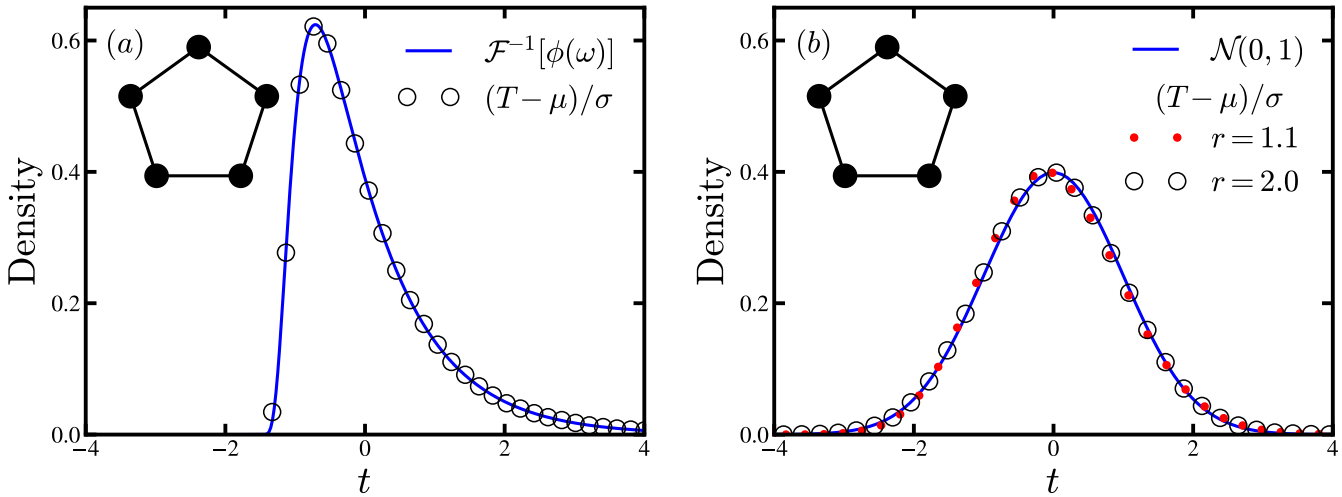

FIG. 1. Fixation-time distributions on the 1D lattice obtained from 10^6 simulation runs. All distributions are standardized to zero mean and unit variance. Solid curves are the theoretical predictions. Shown are the fixation-time distributions for (a) a 1D lattice of $N = 100$ nodes with neutral fitness $r = 1$ and (b) a 1D lattice of $N = 5000$ nodes with mutant fitnesses $r = 1.1$ and $r = 2.0$. For the neutral fitness case, the theoretical distribution was generated by numerical inverse Fourier transform of the characteristic function (Eq. 12). The $r = 1.1$ distribution is slightly but visibly skewed due to finite network size.

of the random walk. As $N \to \infty$, long walks with large fixation times become common and the system revisits each state infinitely often [44].

Knowledge of the cumulants allows us to obtain the exact characteristic function of the fixation-time distribution:

$$\phi(\omega) = e^{-\sqrt{\frac{5}{2}}\omega}\,\Gamma\left(1 - \frac{90^{1/4}\sqrt{\omega}}{\pi}\right)\Gamma\left(1 + \frac{90^{1/4}\sqrt{\omega}}{\pi}\right). \tag{12}$$

Although we cannot find a simple expression for the distribution itself, we can efficiently evaluate it by taking the inverse Fourier transform of the characteristic function numerically. Figure 1(a) shows that the predicted fixation-time distribution agrees well with numerical experiments.

Non-neutral fitness

Next, consider $r \neq 1$ with the transition probabilities given by Eq. 9. Then the eigenvalues of the transition matrix are no longer expressible in closed form. If r is not too large, however, the probabilities b_m and d_m do not vary dramatically with m, the number of mutants. In particular, $b_m \sim 1/N$ for all m when N is large. Therefore, as a first approximation we treat the Bd process on a 1D lattice as a biased random walk with $b_m = r/(1+r)$ and $d_m = 1/(1+r)$. The eigenvalues of the corresponding transition matrix are

$$\lambda_m = 1 - \frac{2\sqrt{r}}{1+r}\cos\left(\frac{m\pi}{N}\right), \quad m = 1, 2, \ldots, N-1. \tag{13}$$

The cumulants again involve sums $S_n = \sum_{m=1}^{N-1}(\lambda_m)^{-n}$, which can be approximated in the limit $N \to \infty$ by an integral,

$$S_n \approx \frac{N}{\pi}\int_0^{\pi} \frac{1}{(1 - 2\sqrt{r}/(1+r)\cos x)^n}\,\mathrm{d}x. \tag{14}$$

Since the integral is independent of N and converges for $r \neq 1$, each of the sums scales linearly: $S_n \sim N$. Thus, using Eq. 4, we see that all cumulants past second order approach 0,

$$\kappa_n \sim \frac{1}{N^{(n-2)/2}} \xrightarrow{N\to\infty} 0, \quad n \geq 3. \tag{15}$$

Hence the fixation-time distribution is asymptotically normal, independent of fitness level.

By evaluating the integrals in Eq. 14, we can more precisely compute the scaling of the cumulants. For the skew we find

$$\kappa_3 \approx \frac{2 + 2r(r+4)}{(r+1)\sqrt{(r^2-1)}}\frac{1}{\sqrt{N}}. \tag{16}$$

This integral approximation becomes accurate when the first term in the sums S_n for $n = 2$ and 3 becomes close to the value of the integrand evaluated at the lower bound ($x = 0$). The fractional difference between these quantities is

$$\begin{aligned}\Delta &= \left|\frac{(1 - 2\sqrt{r}/(1+r))^n}{(1 - 2\sqrt{r}/(1+r)\cos(\pi/N))^n} - 1\right| \\ &= \frac{\sqrt{r}n\pi^2}{(\sqrt{r}-1)^2 N^2} + \mathcal{O}(1/N^4).\end{aligned} \tag{17}$$

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

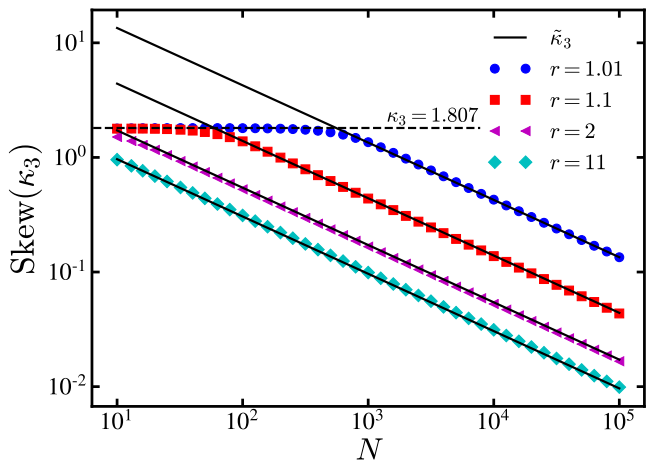

FIG. 2. Scaling of the skew of the fixation-time distribution on the 1D lattice with non-neutral fitness. Data points show numerical calculation of the skew for various fitness levels. The solid lines are the predicted scaling given in Eq. 19 with exponent $q = 1/2$ for each value of fitness r. For small N (and small enough r), the skew is that of a random walk, namely $\kappa_3 = 1.807$, as shown by the dashed line. For large N, the skew $\kappa_3 \sim 1/\sqrt{N}$ with an r-dependent coefficient.

Then we have $\Delta \ll 1$ when $N \gg N_c \approx 2\pi\sqrt{n}/(r-1)$ (assuming r is near 1). To compute the skew, we require the sums with $n = 2$ and 3, giving $N_c \approx 10/(r-1)$.

The above calculation fails for $r \gg 1$, because when $r = \infty$ the transition probabilities $b_m = 1/m$ have different asymptotic behavior as $N \to \infty$. In particular, more time is spent waiting at states with large m. The process still has normally distributed fixation times, but the skew becomes

$$\kappa_3^\infty = 2\left(\sum_{m=1}^{N-1} m^3\right) \Big/ \left(\sum_{m=1}^{N-1} m^2\right)^{3/2} \approx \frac{3\sqrt{3}}{2}\frac{1}{\sqrt{N}}, \quad (18)$$

for large N. Notice that the coefficient is different from that given by the infinite-r limit of Eq. 16, $\kappa_3 \approx 2/\sqrt{N}$. We conjecture that there is a smooth crossover between these two scaling laws with the true skew given approximately by

$$\tilde{\kappa}_3 = \kappa_3\left[r^{-q} + \frac{3\sqrt{3}}{4}(1 - r^{-q})\right] \quad (19)$$

for some exponent q, where κ_3 is the skew given in Eq. 16. For small r this ansatz has skew similar to that of a random walk, but captures the correct large-r limit. We do not have precise theoretical motivation for this ansatz, but as discussed below, it appears to work quite well.

Numerical calculation of the skew for the 1D lattice was performed using the recurrence relation method discussed above and detailed in *SI Abstract*, S3. The results are shown in Fig. 2 for a few values of r. This calculation confirms our initial hypothesis, that near neutral fitness the waiting times are uniform enough that the process is well approximated by a biased random walk and the skew approaches 0, scaling in excellent agreement with Eq. 16. When $N \ll N_c$, the bias is not sufficient to give the mutants a substantial advantage: the process behaves like an unbiased random walk and the fixation-time distribution has large skew $\kappa_3 \approx 1.807$, as found in the preceding section. For $N \gg N_c$, the bias takes over, the cumulants approach 0, and the distribution become normal. For large fitness $r \gg 1$, the crossover ansatz (Eq. 19) captures the scaling behavior quite well if we use an exponent $q = 1/2$. Direct numerical simulations of the process confirm that, for any $r > 1$, the fixation time on the 1D lattice has an asymptotically normal distribution [Fig. 1(b)].

The random walk approximation allows us to find the scaling law for the fixation-time cumulants, but ignores the heterogeneity of waiting times present in the Moran process. Using visit statistics we can compute the cumulants exactly with Eq. 8 and rigorously prove they vanish as $N \to \infty$, verifying that the waiting times have no influence on the asymptotic form of the distribution. Details are provided in *SI Abstract*, S4.

COMPLETE GRAPH

We now turn to the Moran process on a complete graph, useful for modeling well-mixed populations in which all individuals interact. By similar reasoning to above, given m mutants the probability of adding a mutant in the next time step is

$$b_m = \frac{rm}{rm + N - m} \cdot \frac{N - m}{N - 1}, \quad (20)$$

while the probability of subtracting a mutant is

$$d_m = \frac{N - m}{rm + N - m} \cdot \frac{m}{N - 1}. \quad (21)$$

Interestingly, as we will see in this section, these transition probabilities give rise to a rich fitness dependence of the fixation-time distribution in stark contrast to the universality observed on the 1D lattice.

Neutral Fitness

Again we begin with neutral fitness $r = 1$. Now $b_m = d_m = (Nm - m^2)/(N^2 - N)$. The eigenvalues of this transition matrix also have a nice analytical form:

$$\lambda_m = \frac{m(m+1)}{N(N-1)}, \quad m = 1, 2, \ldots N - 1. \quad (22)$$

The asymptotic form of the sums S_n (defined identically to above), can be found by taking the partial fraction

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

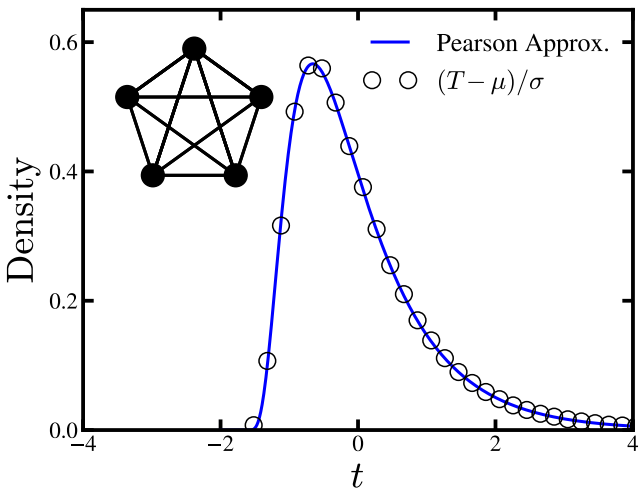

FIG. 3. Fixation-time distributions on the complete graph with $N = 100$ nodes and neutral fitness ($r = 1$) obtained from 10^6 simulation runs. The distribution is standardized to zero mean and unit variance. The solid curve is a fifth order Pearson approximation to the theoretical distribution using the first ten cumulants in Eq. 23.

decomposition of $(\lambda_m)^{-n}$ and evaluating each term individually. The resulting cumulants are

$$\kappa_n = (n-1)!\frac{3^{n/2}}{(\pi^2-9)^{n/2}} \times (-1)^n \sum_{k=1}^{n} \binom{2n-k-1}{n-1} \left[\zeta(k)\left(1-(-1)^k\right)-1\right]. \tag{23}$$

We have not managed to calculate the distribution of the fixation times or its characteristic function in closed form, but fortunately the full set of its cumulants provides equivalent information. By using the first ten cumulants given by Eq. 23, we approximate the true distribution with a fifth-order Pearson-style distribution [as described in Ref. [45], Chapter 5]. Figure 3 shows that the result agrees well with simulations.

The numerical value of the fixation-time skew for the Bd process on the complete graph is $\kappa_3 \approx 1.671$, slightly less than that for the 1D lattice. The decrease is the result of two competing effects contributing to skew. First, since the birth and death transition probabilities are the same, the process is a random walk, which has a highly skewed fixation-time distribution, as shown above. The average time spent in each state, however, varies with m. For instance, when $m = 1$ or $N-1$, $b_m \to 0$ for large N. But if $m = \alpha N$ for some constant $0 < \alpha < 1$ independent of N, then b_m approaches a constant. The beginning and end of the process are very slow because the transition probabilities are vanishingly small. To start, the single mutant must be chosen to give birth from the N available nodes. Near fixation, the reproducing mutant must find and replace one of the few remaining non-mutants.

The characteristic slowing down at certain states is reminiscent of "coupon collection", as discussed earlier. Erdős and Rényi proved that for large N, the normalized time to complete the coupon collection follows a Gumbel distribution [41], which we denote by Gumbel(α, β) with density

$$f(t) = \beta^{-1} e^{-(t-\alpha)/\beta} \exp(-e^{-(t-\alpha)/\beta}). \tag{24}$$

For the Moran process, each slow region is produced by long waits for the random selection of rare types of individuals: either mutants near the beginning of the process or non-mutants near the end. In fact, for infinite fitness there is an exact mapping between Bd on the complete graph and coupon collection [37, 38]. In the next section we show that the two coupon collection regions of the Bd process on a complete graph lead to fixation-time distributions that are convolutions of two Gumbel distributions. In the case of neutral fitness, these Gumbel distributions combine with the random walk to produce a new highly skewed distribution with cumulants given by Eq. 23.

Non-neutral fitness

We saw above that when the average time spent in each state is constant or slowly varying, as is the case for the biased random walk and Moran Bd on the 1D lattice respectively, the fixation-time distribution is normal. In contrast to these systems, Moran Bd on the complete graph exhibits coupon collection regions at the beginning and end of the process, where the transition probabilities are very small. We begin this section with a heuristic argument that correctly gives the asymptotic fixation-time distribution in terms of independent iterations of coupon collection.

Differentiating b_m with respect to m, we find the slope near $m = 0$ is $(r+1)/N$, while the slope near $m = N$ has magnitude $(r+1)/(rN)$ for $N \gg 1$. The transition rates approach zero at each of these points, so we expect behavior similar to coupon collection giving rise to two Gumbel distributions. Since the slope is greater for m near 0 than for m near N, the Moran process completes its coupon collection faster near the beginning of the process than near fixation.

This heuristic suggests that the asymptotic fixation time should be equal in distribution to the sum of two Gumbel random variables, one weighted by r, which is the ratio of the slopes in the coupon collection regions. Specifically, if T is the fixation time with mean μ and variance σ^2, we expect

$$\frac{T-\mu}{\sigma} \xrightarrow{d} \frac{G+rG}{\sqrt{1+r^2}}, \tag{25}$$

where $\xrightarrow{d}$ means convergence in distribution for large N and G is a Gumbel random variable with zero mean and unit variance. It is easy to check that G has distribution

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

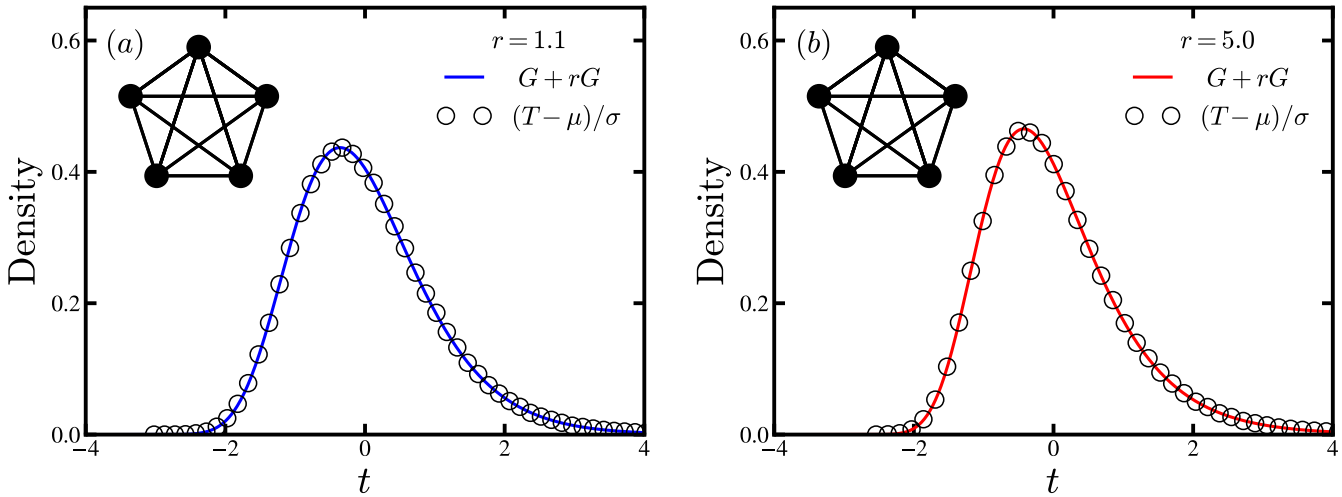

FIG. 4. Fixation-time distributions on the complete graph with $N = 5000$ nodes and non-neutral fitness ($r > 1$) obtained from 10^6 simulation runs. All distributions are standardized to zero mean and unit variance. Solid curves are the theoretical predictions obtained by numerical convolution of two Gumbel distributions, one weighted by r. Distributions are shown for (a) $r = 1.1$ and (b) $r = 5.0$. For larger r, the distribution has larger skew and a slightly sharper peak.

Gumbel$(-\gamma\sqrt{6}/\pi, \sqrt{6}/\pi)$, where $\gamma \approx 0.5772$ is the Euler-Mascheroni constant.

Let us make this argument more rigorous. Previous theoretical analysis showed that in the infinite fitness limit, the fixation time has an asymptotically Gumbel distribution [38]. This result can be recovered within our framework, since when $r = \infty$ it follows that $d_m = 0$, so the eigenvalues of the transition matrix are just $b_m = (N - m)/(N - 1)$ and the cumulants can be directly calculated using Eq. 4.

Now consider large, but not infinite, fitness. In this limit, the number of mutants is monotonically increasing, to good approximation, since the probability that the next change in state increases the mutant population is $r/(1 + r) \approx 1$. The time spent waiting in each state, however, changes dramatically, especially near $m = 1$. Here, $b_1 \to 0$ for large N, in stark contrast to the infinite fitness system where $b_1 \to 1$. The time spent at each state, t_m is an exponential random variable, $\mathcal{E}(b_m + d_m)$. In this approximation each state is visited exactly once, so the total fixation time is a sum of these waiting times:

$$T = \sum_{m=1}^{N-1} \mathcal{E}(b_m + d_m). \tag{26}$$

But this sum of exponential random variables has density given by Eq. 3, with the substitution $\lambda_m \to b_m + d_m$. Thus, the cumulants of $(T - \mu)/\sigma$ are

$$\kappa_n = \frac{1 + r^n}{(1 + r^2)^{n/2}} \cdot \frac{(n-1)!\zeta(n)}{\zeta(2)^{n/2}}, \tag{27}$$

which are exactly the cumulants corresponding to the sum of Gumbel random variables given in Eq. 25. In the limit $r \to \infty$, the first term in Eq. 27 becomes 1, and the cumulants are those for a single Gumbel distribution, in agreement with previous results.

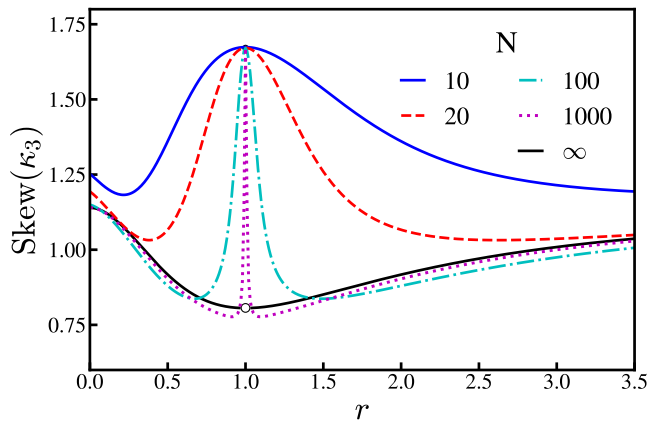

FIG. 5. Fitness dependence of fixation-time skew for the Moran Birth-death process on the complete graph. The skew is shown for $r \geq 0$ and is invariant under $r \to 1/r$. For finite N, the skew does not have a discontinuity, but does show non-monotonic dependence on fitness r. In particular, for a given N, there is a certain fitness level with minimum skew. As $N \to \infty$, we see non-uniform convergence to the predicted skew given by κ_3 in Eq. 27, leading to the discontinuity at $r = 1$. Moreover, for fixed r, the convergence to the $N = \infty$ skew is non-monotonic.

Remarkably, these cumulants are exact for any $r > 1$, not just in the large-r limit. We can see this directly for the skew κ_3 using the visit statistics approach, computing the asymptotic form of Eq. 8 with the complete graph transition probabilities given in Eqs. 20 and 21. Details of the asymptotic analysis are provided in *SI Ab-*

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

stract, S5. Numerical simulations of the Moran process corroborate our theoretical results. As shown in Fig. 4, for $r = 1.1$ and $r = 5$ the agreement between simulated fixation times and the predicted convolution of Gumbel distributions is excellent, at least when N is sufficiently large.

For smaller networks, it is fascinating to see how the results converge to the asymptotic predictions as N grows. Figure 5 shows how the skew of the fixation-time distribution depends on r and N for the complete graph. Note the fixation-time distributions for these systems are invariant under $r \to 1/r$. Therefore we show the skew for all $r > 0$, to emphasize the intriguing behavior near neutral fitness, where $r = 1$. We find that non-uniform convergence of the fixation-time skew leads to the discontinuity predicted at $r = 1$. For finite N, the skew is a non-monotonic function of r and has a minimum value at some fitness $r_{\min}(N)$. Furthermore, at fixed r, the convergence to the $N = \infty$ limit is itself non-monotone. Though beyond the scope of the current study, further investigation of this finite-N behavior would be worth pursuing.

PARTIAL FIXATION TIMES: TRUNCATING COUPON COLLECTION

In many applications, we may be interested in the time to partial fixation of the network. For instance, considering cancer progression [23, 24, 46] or the incubation of infectious diseases [38], symptoms can appear in a patient even when a relatively small proportion of cells are malignant or infected. We therefore consider T_α, the total time to first reach αN mutants on the network, where $0 < \alpha < 1$. The methods developed above apply to these processes as well. For the eigen-decomposition approach we instead use the sub-matrix of Ω_{tr} containing the first αN rows and columns. In calculations involving the numerical recurrence relations or visit statistics, we simply cut the sums off at αN instead of N and for the latter, replace $w_{i_1 i_2 \cdots i_n}(r, N)$ with $w_{i_1 i_2 \cdots i_n}(r, \alpha N)$.

Truncation of the Moran Bd process on the 1D lattice has no effect on the asymptotic shape of the fixation-time distributions. In both the neutral fitness system and the random walk approximation to the non-neutral fitness system, the transition matrix has no explicit dependence on the state or N (aside from proportionality factors that cancel in Eq. 4). Thus, the eigenvalues are identical to those calculated previously, but correspond to a smaller effective system size αN. Taking the limit $N \to \infty$ therefore yields the same asymptotic distribution found above. This result is borne out numerically.

In contrast to the 1D lattice, the complete graph exhibits more interesting dependence on truncation. Since the transition probabilities have state dependence, the eigenvalues vary with truncation (they don't exactly correspond to the full system with a smaller effective N). Our intuition from coupon collection, however, lets us predict the resulting distribution.

First suppose fitness is non-neutral. Then there are two coupon collection stages, one near the beginning and another near the end of the process, and together they generate a fixation-time distribution that is a weighted convolution of two Gumbel distributions. The effect of truncating the process near its end should now become clear: it simply removes the second coupon collection, since truncation means the process stops before the mutants have to laboriously find and replace the last remaining non-mutants. Therefore, we intuitively expect the fixation time for non-neutral fitness to be distributed according to just a single Gumbel distribution, regardless of fitness level.

The only exception occurs if $r = \infty$; then no coupon collection occurs at the beginning of the process either, because the lone mutant is guaranteed to be selected to give birth in the first time step, thanks to its infinite fitness advantage. Thus, when fitness is infinite and the process is truncated at the end, both coupon collection phases are deleted. In this case, the fixation times are normally distributed.

Similar reasoning applies to the Birth-death process with neutral fitness. It also has two coupon collection regions, one of which is removed by truncation. In this case, however, the random walk mechanism contributes to the skew of the overall fixation-time distribution, combining non-trivially with the coupon collection-like process. We find that the skew of the fixation time depends on the truncation factor α, varying between $6\sqrt{3}(10 - \pi^2)/(\pi^2 - 9)^{3/2} \approx 1.6711$ when $\alpha = 1$, and $\sqrt{3} \approx 1.732$ when $\alpha = 0$. A derivation of the $\alpha \to 0$ limit of the skew is given in *SI Abstract*, S5.

The above results are summarized in Fig. 6, which shows the asymptotic fitness dependence of fixation-time skew for each network considered in this paper. As before, we show the skew for all $r > 0$ (not just $r > 1$) to emphasize the discontinuities at zero, neutral, and infinite fitness r. On the 1D lattice, independent of the truncation factor α, the Bd process has normally distributed fixation times, except at neutral fitness where the distribution is highly skewed. The complete graph fixation-time distributions are the weighted convolution of two Gumbel distributions for $r \neq 1$, again with a highly skewed distribution at $r = 1$. With truncation by a factor α, the distribution for the complete graph is a Gumbel for $1 < r < \infty$, and a normal for $r = \infty$. With neutral fitness the fixation distribution is again highly skewed, with skew dependent on the truncation factor α.

DISCUSSION

Multi-fitness Moran process: comparison to the Isotherm Theorem

So far, we have considered Moran processes with a single fitness level, designating the relative reproduction

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

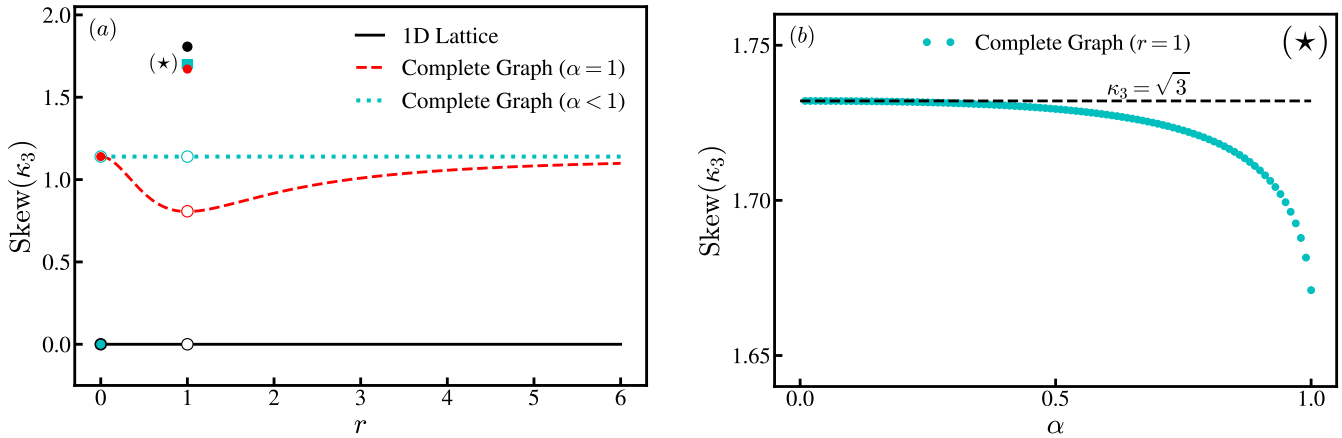

FIG. 6. Variation of fixation-time skew with fitness level r and truncation factor α for different network structures. (a) The fixation-time skew versus fitness for the 1D lattice (black solid line), complete graph (red dashed line), and complete graph with truncation (green dotted line). The skew is shown for all $r \geq 0$ and is invariant under $r \to 1/r$. Each curve has a discontinuity at $r = 1$, where the distribution is highly skewed and jumps up to values of $\kappa_3 > 1.5$. The truncated fixation time on the complete graph has a second discontinuity at $r = \infty$ (equivalently $r = 0$), where the distribution goes from Gumbel to normal. The complete graph with truncation also has variable skew at $r = 1$, dependent on the value of α, shown by the solid bar labelled with ($\star$). (b) The fixation-time skew for the complete graph with neutral fitness, plotted as a function of α.

rates between mutants and non-mutants. It also makes sense to consider a second fitness level $\tilde{r}$ during the replacement step, measuring the resilience of mutants versus non-mutants [9]. Taking this into account, when a mutant or non-mutant is trying to replace its neighbors, mutants are replaced with probability proportional to $1/\tilde{r}$. Taking $\tilde{r} = 1$ returns to the model used throughout the preceding sections. The multi-fitness model may better capture the complexity of real-world evolutionary systems but, as we will see below, does not generally give rise to qualitatively different fixation-time distributions.

On the 1D lattice, the Moran process with fitness at both steps (birth and death), has new transition probabilities

$$b_m = \frac{r}{rm + N - m}\frac{\tilde{r}}{1+\tilde{r}}, \quad d_m = \frac{1}{r\tilde{r}}b_m, \tag{28}$$

for $1 < m < N - 1$. The probabilities are different when $m = 1$ and $m = N - 1$ and there is only one mutant or non-mutant. In this case the cells on the population boundary don't have one mutant and one non-mutant as neighbors, as is the case for all other m. In the limit $N \gg 1$, however, changing these two probabilities does not affect the fixation-time distribution and we can use the probabilities given in Eq. 28.

The multi-fitness Moran model on the 1D lattice differs from the previously considered Moran Bd process in two ways. First, the transition probabilities have the same functional form as before, but are scaled by a factor $\tilde{r}(1+\tilde{r})^{-1}$. This factor determines the time-scale of the process but does not alter the shape of the fixation-time distribution because it drops out of the expression for the cumulants, Eq. 4. Second, the ratio $b_m/d_m = r\tilde{r}$ shows that the process is still a random walk, but with new bias corresponding to an effective fitness level $r_{\text{eff}} = r\tilde{r}$. With these observations, when $r_{\text{eff}} \neq 1$, our preceding analysis applies and we predict normally distributed fixation times. If $r_{\text{eff}} = 1$, the random walk is unbiased, and we expect highly skewed fixation-time distributions.

On the complete graph, considering fitness during the replacement step leads to transition probabilities

$$b_m = \frac{rm}{rm + N - m} \cdot \frac{\tilde{r}(N-m)}{\tilde{r}(N-m) + m - 1} \tag{29}$$

and

$$d_m = \frac{N-m}{rm + N - m} \cdot \frac{m}{\tilde{r}(N-m-1) + m}. \tag{30}$$

In this case, the ratio of transition probabilities is m-dependent, but as $N \to \infty$, $b_m/d_m \to r\tilde{r}$, again motivating the definition of the effective fitness level $r_{\text{eff}} = r\tilde{r}$. If we take the large (but not infinite) fitness limit $r_{\text{eff}} \gg 1$, so that the mutant population is monotonically increasing to good approximation, then the fixation time cumulants are again given by Eq. 4 with $\lambda_m \to b_m + d_m$. As $N \to \infty$, the cumulants become

$$\kappa_n = \frac{1 + r^n/\tilde{r}^n}{(1 + r^2/\tilde{r}^2)^{n/2}} \cdot \frac{(n-1)!\zeta(n)}{\zeta(2)^{n/2}}, \tag{31}$$

identical to the Moran Bd process on the complete graph, with $r \to r/\tilde{r}$. Simulations again indicate this expression for the cumulants holds for all r, not just in the high fitness limit. When $r_{\text{eff}} = 1$, we expect highly skewed fixation distributions arising from the unbiased random walk underlying the dynamics. This is indeed the case, though simulations show we get a whole family of distributions dependent on $r = 1/\tilde{r}$.

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

It is interesting to contrast the above observations with a result in evolutionary dynamics known as the isotherm theorem. The theorem states that for $\tilde{r} = 1$, the Moran process on a large class of networks, known as isothermal graphs, has fixation probability identical to the complete graph [5]. Recent work has shown that this breaks down if $\tilde{r} \neq 1$; the fixation probability develops new network dependence [9]. In contrast, even isothermal graphs (including the complete graph and 1D lattice) have fixation-time distributions that depend on network structure. Likewise, adding a second fitness level to the model breaks the universality in fixation probabilities predicted by the isotherm theorem, but leads to the same family of fixation distributions that arise due to network structure.

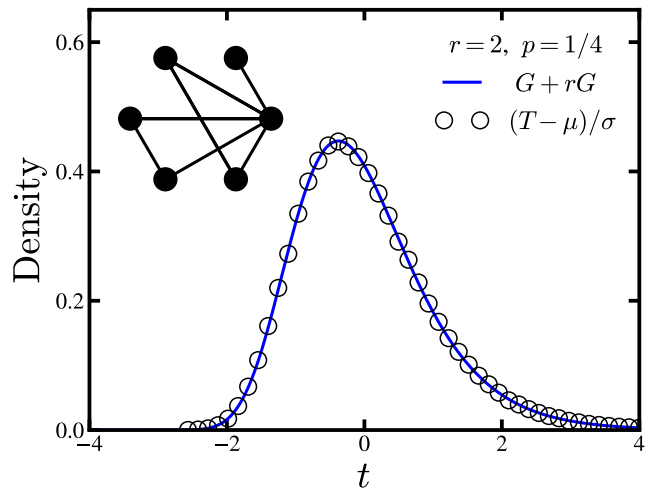

FIG. 7. Fixation-time distribution on an Erdős-Rényi random graph with $N = 100$ nodes, edge probability $p = 1/4$, and fitness $r = 2$, obtained from 10^6 simulation runs. The distribution is standardized to zero mean and unit variance. The solid curve is the theoretical prediction for the complete graph, obtained by numerical convolution of two Gumbel distributions, one weighted by r. For these parameters, the random graph fixation time is accurately captured by the mean field approximation.

Mean-Field Transition Probabilities: Random Networks, Stars, and Lattices

While the 1D lattice and complete graph provide illustrative exactly solvable models of the fitness dependence of fixation-time distributions, other networks may be more realistic. On more complicated networks many of the analytical tools developed here fail because the transition probabilities (the probability of adding or subtracting a mutant given the current state) depend on the full configuration of mutants, not just the number of mutants. In principle such a system could still be modeled as a Markov process, but the state space would become prohibitively large.

Fortunately, for certain networks the effect of different configurations can be averaged over, giving a mean-field approximation to the transition probabilities. This approach has been used on a variety of networks to calculate fixation times with reasonable success [31, 35]. For instance, the mean-field transition probabilities for the Erdős-Rényi random network were recently estimated in Ref. [31]. The result is identical to the complete graph probabilities (Eqs. 20-21) up to a constant factor $1 - 2/Np$, dependent on the edge probability p for the network. This correction is important for computing the mean fixation times, but does not affect the shape of the fixation-time distribution, since proportionality factors cancel in the expression for the cumulants. Therefore we again expect a fixation-time distribution corresponding to the weighted sum of two Gumbel distributions. Previous work indicates this prediction holds for infinite fitness, where the fixation time on an Erdős-Rényi network has a Gumbel distribution [38].

Preliminary simulations show that the Erdős-Rényi network has the expected fixation-time distributions for $p = 1/4$ and $r = 2$ (see Figure 7). Whether this result holds for the full range of fitness and edge probabilities requires further investigation. It may be that for some values of p and r the mean-field approximation is not sufficient to capture the higher-order moments of the distribution. In the same vein, what other networks admit accurate mean-field approximations to the transition probabilities? Do many complex networks have fixation-time distributions identical to the complete graph?

Another nice approximation maps the Moran process on a star graph, a simple amplifier of selection, onto a birth-death Markov chain [15]. The resulting transition probabilities exhibit coupon collection regions, similar to the complete graph. The ratio of slopes near these regions (few mutants or non-mutants), however, is r^2. Our heuristic predicts the fixation-time distribution on the star is $G + r^2 G$. In addition to amplifying fixation probability, the star increases fixation-time skew. This raises a broader question: do evolutionary amplifiers also amplify fixation-time skew?

Similar mean-field arguments have also been applied to d-dimensional lattices in the infinite-fitness limit [37, 38]. In this limit the population of mutants grows in an approximately spherical shape near the beginning of the process and the population of non-mutants is approximately spherical near fixation. The surface area to volume ratio of the d-dimensional sphere gives the probability of adding a mutant. With finite fitness non-mutants can now replace their counterparts and the surface of the sphere of growing mutants roughens [46]. For near-neutral fitness, the configurations of mutants resembles the shape of real cancerous tumors. Perhaps, by extending our theoretical framework, connections can be made between the fitness-dependent roughness of a growing mutant population and fixation-time distributions for Moran processes on lattices.

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Future Directions

Future studies could analyze random networks and lattices more deeply, as well as stars, the prototypical evolutionary amplifiers [5]. More sophisticated models involving evolutionary games are also of interest. These have skewed fixation-time distributions [34] whose asymptotic form remains unknown. Finally, we hope that methods developed here will prove useful in other areas, such as epidemiology [47], ecology [48], and protein folding [49], where stochastic dynamics may similarly give rise to skewed first-passage times.

ACKNOWLEDGMENTS

Research supported by an NSF Graduate Research fellowship grant DGE-1650441 to D. H. and by NSF grants DMS-1513179 and CCF-1522054 to S.H.S.

[1] Nowak MA (2006) *Evolutionary dynamics*. (Harvard University Press).

[2] Maruyama T (1974) A markov process of gene frequency change in a geographically structured population. *Genetics* 76(2):367–377.

[3] Maruyama T (1974) A simple proof that certain quantities are independent of the geographical structure of population. *Theoretical population biology* 5(2):148–154.

[4] Slatkin M (1981) Fixation probabilities and fixation times in a subdivided population. *Evolution* 35(3):477–488.

[5] Lieberman E, Hauert C, Nowak MA (2005) Evolutionary dynamics on graphs. *Nature* 433(7023):312.

[6] Antal T, Scheuring I (2006) Fixation of strategies for an evolutionary game in finite populations. *Bulletin of mathematical biology* 68(8):1923–1944.

[7] Houchmandzadeh B, Vallade M (2011) The fixation probability of a beneficial mutation in a geographically structured population. *New Journal of Physics* 13(7):073020.

[8] Díaz J, et al. (2014) Approximating fixation probabilities in the generalized moran process. *Algorithmica* 69(1):78–91.

[9] Kaveh K, Komarova NL, Kohandel M (2015) The duality of spatial death–birth and birth–death processes and limitations of the isothermal theorem. *Royal Society open science* 2(4):140465.

[10] Jamieson-Lane A, Hauert C (2015) Fixation probabilities on superstars, revisited and revised. *Journal of Theoretical Biology* 382:44 – 56.

[11] Altrock PM, Traulsen A, Nowak MA (2017) Evolutionary games on cycles with strong selection. *Physical Review E* 95(2):022407.

[12] Tkadlec J, Pavlogiannis A, Chatterjee K, Nowak MA (2018) Fixation probability and fixation time in structured populations. *arXiv preprint arXiv:1810.02687*.

[13] Hindersin L, Traulsen A (2015) Most undirected random graphs are amplifiers of selection for birth-death dynamics, but suppressors of selection for death-birth dynamics. *PLOS Computational Biology* 11(11):1–14.

[14] Goldberg LA, et al. (2018) Asymptotically optimal amplifiers for the moran process. *Theoretical Computer Science*.

[15] Frean M, Rainey PB, Traulsen A (2013) The effect of population structure on the rate of evolution. *Proceedings of the Royal Society of London B: Biological Sciences* 280(1762).

[16] Sella G, Hirsh AE (2005) The application of statistical physics to evolutionary biology. *Proceedings of the National Academy of Sciences* 102(27):9541–9546.

[17] Barton N, Coe J (2009) On the application of statistical physics to evolutionary biology. *Journal of theoretical biology* 259(2):317–324.

[18] Pavlogiannis A, Tkadlec J, Chatterjee K, Nowak MA (2018) Construction of arbitrarily strong amplifiers of natural selection using evolutionary graph theory. *Communications Biology* 1(1):71.

[19] Nowak MA, Sigmund K (2004) Evolutionary dynamics of biological games. *science* 303(5659):793–799.

[20] Altrock PM, Traulsen A (2009) Fixation times in evolutionary games under weak selection. *New Journal of Physics* 11(1):013012.

[21] Gerlee P, Altrock PM (2017) Extinction rates in tumour public goods games. *Journal of The Royal Society Interface* 14(134):20170342.

[22] Allen B, et al. (2017) Evolutionary dynamics on any population structure. *Nature* 544(7649):227.

[23] Sottoriva A, et al. (2013) Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proceedings of the National Academy of Sciences* 110(10):4009–4014.

[24] Bozic I, et al. (2013) Evolutionary dynamics of cancer in response to targeted combination therapy. *elife* 2:e00747.

[25] Noble C, Olejarz J, Esvelt KM, Church GM, Nowak MA (2017) Evolutionary dynamics of crispr gene drives. *Science Advances* 3(4):e1601964.

[26] Smith RJ, Okano JT, Kahn JS, Bodine EN, Blower S (2010) Evolutionary dynamics of complex networks of hiv drug-resistant strains: the case of san francisco. *Science* 327(5966):697–701.

[27] Kimura M (1980) Average time until fixation of a mutant allele in a finite population under continued mutation pressure: Studies by analytical, numerical, and pseudo-sampling methods. *Proceedings of the National Academy of Sciences* 77(1):522–526.

[28] Askari M, Samani KA (2015) Analytical calculation of average fixation time in evolutionary graphs. *Physical Review E* 92(4):042707.

[29] Askari M, Miraghaei ZM, Samani KA (2017) The effect of hubs and shortcuts on fixation time in evolutionary graphs. *Journal of Statistical Mechanics: Theory and Experiment* 2017(7):073501.

[30] Farhang-Sardroodi S, Darooneh AH, Nikbakht M, Komarova NL, Kohandel M (2017) The effect of spatial randomness on the average fixation time of mutants. *PLoS computational biology* 13(11):e1005864.

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[31] Hajihashemi M, Aghababaei Samani K (2018) Fixation time in evolutionary graphs: a mean field approach. *bioRxiv*.

[32] Dingli D, Traulsen A, Pacheco JM (2007) Stochastic dynamics of hematopoietic tumor stem cells. *Cell Cycle* 6(4):461–466.

[33] Altrock PM, Traulsen A, Reed FA (2011) Stability properties of underdominance in finite subdivided populations. *PLoS computational biology* 7(11):e1002260.

[34] Ashcroft P, Traulsen A, Galla T (2015) When the mean is not enough: Calculating fixation time distributions in birth-death processes. *Physical Review E* 92(4):042154.

[35] Ying LM, Zhou J, Tang M, Guan SG, Zou Y (2018) Mean-field approximations of fixation time distributions of evolutionary game dynamics on graphs. *Frontiers of Physics* 13(1):130201.

[36] Aldous D, , et al. (2013) Interacting particle systems as stochastic social dynamics. *Bernoulli* 19(4):1122–1149.

[37] Ottino-Löffler B, Scott JG, Strogatz SH (2017) Takeover times for a simple model of network infection. *Physical Review E* 96(1):012313.

[38] Ottino-Löffler B, Scott JG, Strogatz SH (2017) Evolutionary dynamics of incubation periods. *eLife* 6:e30212.

[39] Moran PAP (1958) The effect of selection in a haploid genetic population in *Mathematical Proceedings of the Cambridge Philosophical Society*. (Cambridge University Press), Vol. 54, pp. 463–467.

[40] Moran PAP (1958) Random processes in genetics in *Mathematical Proceedings of the Cambridge Philosophical Society*. (Cambridge University Press), Vol. 54, pp. 60–71.

[41] Erdős P, Rényi A (1961) On a classical problem of probability theory. *Publ. Math. Inst. Hung. Acad. Sci.* 6:215–220.

[42] Bakhtin Y (2018) Universal statistics of incubation periods and other detection times via diffusion models. *arXiv preprint arXiv:1804.05961*.

[43] Keilson J (1965) A review of transient behavior in regular diffusion and birth-death processes. part ii. *Journal of Applied Probability* 2(2):405–428.

[44] Durrett R (2010) *Probability: theory and examples*. (Cambridge university press).

[45] Rose C, Smith M (2002) *Mathematical Statistic with Mathematica*. (Springer-Verlag).

[46] Williams T, Bjerknes R (1972) Stochastic model for abnormal clone spread through epithelial basal layer. *Nature* 236(5340):19.

[47] Doering CR, Sargsyan KV, Sander LM, Vanden-Eijnden E (2007) Asymptotics of rare events in birth–death processes bypassing the exact solutions. *Journal of Physics: Condensed Matter* 19(6):065145.

[48] Doering CR, Sargsyan KV, Sander LM (2005) Extinction times for birth-death processes: Exact results, continuum asymptotics, and the failure of the fokker–planck approximation. *Multiscale Modeling & Simulation* 3(2):283–299.

[49] Zwanzig R, Szabo A, Bagchi B (1992) Levinthal's paradox. *Proceedings of the National Academy of Sciences* 89(1):20–22.

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Information for "Fitness dependence of the fixation-time distribution for evolutionary dynamics on graphs"

David Hathcock and Steven H. Strogatz

This SI is devoted to providing the mathematical details supporting the results quoted in the main text. We start with the general theory for birth-death Markov chains. In Section S1 we show how to condition the transition probabilities on fixation, producing a Markov chain with identical statistics that is guaranteed to reach fixation. This result is used both in formulating the visit statistics approach (Section S2, which gives an exact series expression for the fixation-time cumulants), and to derive a numerically efficient recurrence relation for calculating the fixation-time skew (Section S3). In Sections S4 and S5 we apply these results to the Moran process on the one-dimensional (1D) lattice and complete graph respectively. For the 1D lattice, we compute the asymptotic form of the fixation-time cumulants for neutral fitness and prove the cumulants vanish for non-neutral fitness. For the complete graph, we show the fixation-time skew under non-neutral fitness corresponds to that of a weighted convolution of Gumbel distributions and derive the $\alpha \to 0$ limit of the truncated fixation-time skew in the Moran process with neutral fitness.

S1: BIRTH-DEATH MARKOV CHAIN CONDITIONED ON FIXATION

For both the numerical recurrence relation and the visit statistics approach described in the main text (and detailed below), it is useful to consider the birth-death Markov chain conditioned on hitting N, which has an identical fixation-time distribution. This Markov chain has new conditioned transition probabilities $\tilde{b}_m$ and $\tilde{d}_m$. If X_t is the state of the system at time t, then $\tilde{b}_m = \mathcal{P}(X_t = m \to X_{t+1} = m+1 | X_\infty = N)$ and $\tilde{d}_m$ is defined similarly. Applying the laws of conditional probability, we find that

$$\begin{aligned}\tilde{b}_m &= \frac{\mathcal{P}(X_{t+1} = m+1 \text{ AND } X_t = m \text{ AND } X_\infty = N)}{\mathcal{P}(X_t = m \text{ AND } X_\infty = N)} \\ &= \frac{\mathcal{P}(X_\infty = N | X_t = m+1)}{\mathcal{P}(X_\infty = N | X_t = m)} \mathcal{P}(X_{t+1} = m+1 | X_t = m) \\ &= \frac{\mathcal{P}(X_\infty = N | X_t = m+1)}{\mathcal{P}(X_\infty = N | X_t = m)} b_m,\end{aligned} \tag{1}$$

where b_m is the transition rate in the original Markov chain. Following the same procedure, we also find the backward transition probabilities are related by

$$\tilde{d}_m = \frac{\mathcal{P}(X_\infty = N | X_t = m-1)}{\mathcal{P}(X_\infty = N | X_t = m)} d_m. \tag{2}$$

The conditioned Markov chain has a few nice properties. First, the fixation probability in the conditioned system is one, by construction. This is particularly helpful for accelerating simulations of the Moran process. Conditioning the transition probabilities also accounts for the normalization of the fixation-time distribution. Furthermore, this operation only changes the relative probability of adding versus subtracting a mutant. The probability that the system leaves a given state is unchanged:

$$\begin{aligned}\tilde{b}_m + \tilde{d}_m &= \frac{\mathcal{P}(X_{t+1} = m+1 \text{ AND } X_t = m \text{ AND } X_\infty = N) + \mathcal{P}(X_{t+1} = m-1 \text{ AND } X_t = m \text{ AND } X_\infty = N)}{\mathcal{P}(X_t = m \text{ AND } X_\infty = N)} \\ &= 1 - \frac{\mathcal{P}(X_{t+1} = m \text{ AND } X_t = m \text{ AND } X_\infty = N)}{\mathcal{P}(X_t = m \text{ AND } X_\infty = N)} \\ &= 1 - \mathcal{P}(X_{t+1} = m | X_t = m) \\ &= b_m + d_m.\end{aligned} \tag{3}$$

This invariance, along with Eqs. 1 and 2, shows that conditioning the Markov chain is equivalent to a similarity transformation on the transient transition matrix with a diagonal change of basis:

$$\tilde{\Omega}_{\text{tr}} = S\,\Omega_{\text{tr}}\,S^{-1} \qquad S_{mn} = \mathcal{P}(X_\infty = N | X_t = m)\delta_{m,n}, \tag{4}$$

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

where $\Omega_{\rm tr}$ is the birth-death transition matrix with absorbing states removed as defined in the main text (and in Eq. 7 below).

For the Moran Birth-death process considered in the main text, $b_m/d_m = r$. In this case, by setting up a linear recurrence it is easy to show that the probability of fixation, starting from m mutants, is

$$\mathcal{P}(X_\infty = N|X_t = m) = \frac{1-1/r^m}{1-1/r^N}, \tag{5}$$

so that

$$\tilde{b}_m = \frac{r^{m+1}-1}{r^{m+1}-r} b_m, \qquad \tilde{d}_m = \frac{r^m - r}{r^m - 1} d_m. \tag{6}$$

Note that we can scale the similarity matrix S by an overall constant, so it is convenient to choose $S_{mn} = (1 - 1/r^m)\delta_{m,n}$. For the multi-fitness Moran Birth-Death model discussed in the main text fixation probabilities derived by Kaveh et al. [1] can be used together with Eq. 4 to condition the Markov chain on fixation.

S2: VISIT STATISTICS

In this section we derive an exact series expression for the cumulants of the fixation time. This result requires constant selection, $b_m/d_m = r$, as is the case for the Moran process. In terms of the probabilities b_m and d_m of adding and removing mutants, the transient transition matrix is

$$[\Omega_{\rm tr}]_{mn} = b_n \delta_{m,n+1} + d_n \delta_{m,n-1} - (b_n + d_n)\delta_{m,n}, \tag{7}$$

for $m, n = 1, 2, \ldots, N-1$. As noted in the main text, when the constant selection condition holds this can be written as $\Omega_{\rm RW} D$, where D is diagonal with elements $D_{ii} = b_i + d_i$ and $\Omega_{\rm RW}$ is the transition matrix for a random walk:

$$[\Omega_{\rm RW}]_{nm} = \frac{r}{1+r}\delta_{m,n+1} + \frac{1}{1+r}\delta_{m,n-1} - \delta_{m,n}. \tag{8}$$

Since we are interested in the fixation-time distribution, we want to condition on fixation occurring. From the preceding section, we have that the conditioned transition matrix $\tilde{\Omega}_{\rm tr} = S\,\Omega_{\rm tr}\,S^{-1}$, where S is diagonal with $S_{ii} = 1 - 1/r^i$. Combining these results, we have that

$$\tilde{\Omega}_{\rm tr} = S\,\Omega_{\rm RW}\,S^{-1} D, \tag{9}$$

where we have used the fact that both D and S are diagonal matrices, and therefore commute.

As discussed in the main text, using the conditioned transient transition matrix $\tilde{\Omega}_{\rm tr}$, the fixation-time distribution can be written as

$$f_T(t) = -\mathbf{1}\,\tilde{\Omega}_{\rm tr}\exp(\tilde{\Omega}_{\rm tr} t)\,\mathbf{p}(0), \tag{10}$$

where $\mathbf{1}$ is a row vector of ones and $\mathbf{p}(0)$ is the initial state of the system. In what follows we will always take the initial state to be a single mutant $p_m(0) = \delta_{m,1}$, but generalizations to other cases are straightforward. The corresponding characteristic function is

$$\phi(\omega) := E[\exp(i\omega T)] = \mathbf{1}\,\tilde{\Omega}_{\rm tr}\,(i\omega + \tilde{\Omega}_{\rm tr})^{-1}\,\mathbf{p}(0) \tag{11}$$

and the derivatives $(-i)^n \phi^{(n)}(0)$ give the moments of T,

$$m_n := E[T^n] = (-1)^n n!\,\mathbf{1}\,\tilde{\Omega}_{\rm tr}^{-n}\,\mathbf{p}(0). \tag{12}$$

To compute these moments, we need to compute $\tilde{\Omega}_{\rm tr}^{-1} = D^{-1} S\,\Omega_{RW}^{-1} S^{-1}$. Since Ω_{RW} is a tridiagonal Toeplitz matrix, its inverse has a well-known form [2]:

$$(-\Omega_{RW})^{-1}_{ij} = \begin{cases} \dfrac{(r+1)(r^i-1)(r^N-r^j)}{r^j\,(r-1)(r^N-1)} & \text{if } i \le j, \\ \dfrac{(r+1)(r^j-1)(r^N-r^i)}{r^j\,(r-1)(r^N-1)} & \text{if } i > j. \end{cases} \tag{13}$$

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Hence the matrix $V = S\,\Omega_{\rm RW}^{-1} S^{-1}$ has elements

$$V_{ij} = \begin{cases} \dfrac{(r+1)(r^i-1)^2(r^N-r^j)}{r^i\,(r-1)(r^j-1)(r^N-1)} & \text{if } i \le j, \\ \dfrac{(r+1)(r^i-1)(r^N-r^i)}{r^i\,(r-1)(r^N-1)} & \text{if } i > j. \end{cases} \tag{14}$$

The matrix V, sometimes called the fundamental matrix, encodes the visit statistics of the conditioned random walk: V_{ij} is the mean number of visits to state i from state j before hitting the absorbing state N [3]. The Moran process has the same visit statistics, but on average spends a different amount of time, designated by $b_i + d_i$, waiting in each state.

While one could now compute the moments m_n in Eq. 12 directly, we have found that the cumulants yield nicer expressions. Furthermore, the normal and Gumbel fixation-time distributions, predicted by our simulations and approximate calculations in the main text, are more simply described in terms of their cumulants. The non-standardized cumulants κ'_n are linear combinations involving products of moments whose orders sum to n. Thus each term in the cumulants has n powers of D producing n factors of $b_i + d_i$ with a weight designated by the visit statistics. With this observation, it is clear the standardized cumulants $\kappa_n = \kappa'_n/(\kappa'_2)^{n/2}$ have the form given in the main text,

$$\kappa_n(N) = \left(\sum_{i_1,i_2,\ldots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})} \right) \Bigg/ \left(\sum_{i,j=1}^{N-1} \frac{w^2_{ij}(r,N)}{(b_i+d_i)(b_j+d_j)} \right)^{n/2}, \tag{15}$$

where $w^n_{i_1 i_2 \cdots i_n}(r,N)$ are the weighting factors coming entirely from the visit statistics of a biased random walk. We emphasize that even without explicit knowledge of the factors $w^n_{i_1 i_2 \cdots i_n}(r,N)$, this formulation can be extremely useful. For instance when $b_i + d_i = 1$, these are just the cumulants for the biased random walk, which were computed in the main text to approximate the Moran process on the 1D lattice. This fact can be used to bound the sums in Eq.15, which in some cases is sufficient to determine the leading asymptotic behavior. When this is not possible, the weighting factors must be computed. We derive $w^2_{ij}(r,N)$ and $w^3_{ijk}(r,N)$ below.

We can compute the weighting factors by writing out the matrix multiplication of $\tilde{\Omega}_{\rm tr}^{-1}$. First note that

$$[-\tilde{\Omega}_{\rm tr}^{-1}]_{ij} = \frac{V_{ij}}{b_i + d_i}. \tag{16}$$

Then the first three moments of the fixation time are,

$$\begin{aligned} m_1 &= \sum_{i=1}^{N} \frac{V_{i1}}{b_i+d_i} \\ m_2 &= 2\sum_{i,j=1}^{N} \frac{V_{ij}V_{j1}}{(b_i+d_i)(b_j+d_j)} \\ m_3 &= 6\sum_{i,j,k=1}^{N} \frac{V_{ij}V_{jk}V_{k1}}{(b_i+d_i)(b_j+d_j)(b_k+d_k)}. \end{aligned} \tag{17}$$

The corresponding non-standardized cumulants are given by the usual formulas, $\kappa'_2 = m_2 - m_1^2$ and $\kappa'_3 = m_3 - 3m_2m_1 + 2\mu_1^3$. In terms of the visit numbers the non-standardized cumulants become

$$\begin{aligned} \kappa'_2 &= 2\sum_{i,j=1}^{N} \frac{V_{ij}V_{j1} - V_{i1}V_{j1}}{(b_i+d_i)(b_j+d_j)} \\ \kappa'_3 &= 6\sum_{i,j=1}^{N} \frac{V_{ij}V_{jk}V_{k1} - 3V_{ij}V_{j1}V_{k1} + 2V_{i1}V_{j1}V_{k1}}{(b_i+d_i)(b_j+d_j)(b_k+d_k)}. \end{aligned} \tag{18}$$

From here we can read off the weighting factors accordingly. For convenience, we can choose $w^2_{ij}(r,N)$ and $w^3_{ijk}(r,N)$ to be symmetric by averaging the numerators in Eq. 18 over the permutations of the indices. Then,

$$\begin{aligned} w^2_{ij}(r,N) &= \sum_{\sigma\in\Pi_2} V_{\sigma(1)\sigma(2)}V_{\sigma(2)1} - V_{\sigma(1)1}V_{\sigma(2)1} \\ w^3_{ijk}(r,N) &= \sum_{\sigma\in\Pi_3} V_{\sigma(1)\sigma(2)}V_{\sigma(2)\sigma(3)}V_{\sigma(3)1} - 3V_{\sigma(1)\sigma(2)}V_{\sigma(2)1}V_{\sigma(3)1} + 2V_{\sigma(1)1}V_{\sigma(2)1}V_{\sigma(3)1}, \end{aligned} \tag{19}$$

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

where Π_2 is the set of permutations of $\{i, j\}$ and Π_3 are the permutations of $\{i, j, k\}$. Plugging Eq. 14 into the expression for w_{ij}^2 we obtain, after some algebra,

$$w_{ij}^2(r, N) = \frac{(r+1)^2(r^j-1)^2(r^N-r^i)^2}{r^{i+j}\,(r-1)^2(r^N-1)^2}, \tag{20}$$

for $i \geq j$. Since we have constructed $w_{ij}^2(r, N)$ to be symmetric, when $j > i$ the formula is identical with i and j exchanged. Similarly, using Eq. 14 together with the expression for w_{ijk}^3 in Eq. 19 leads to

$$w_{ijk}^3(r, N) = 2\frac{(r+1)^3(r^k-1)^2(r^j-1)(r^N-r^i)^2(r^N-r^j)}{r^{i+j+k}\,(r-1)^3(r^N-1)^3}, \tag{21}$$

for $i \geq j \geq k$. Again, the formula for different orderings of the indices i, j, k is the same with the indices permuted appropriately, so that w_{ijk}^3 is perfectly symmetric.

This completes the derivation of the visit statistics expression for the fixation-time cumulants. Together, Eqs. 15, 20 and 21 give a closed form expression for the fixation-time skew which is manageable for the purpose of asymptotic approximations. The diagonal terms in the higher-order weighting factors are also particularly simple, $w_{ii\cdots i}^n(r, N) = (n-1)!V_{ii}^n$. While we will not explicitly compute them, the off diagonal weights $w_{i_1 i_2 \cdots i_n}^n(r, N)$ can be found by a straightforward generalization of the above procedure. Applications of this approach are given below, where we show that all cumulants of the fixation time vanish for the Moran process on the 1D lattice and compute the asymptotic skew for the Moran process on the complete graph.

S3: RECURRENCE RELATION FOR FIXATION-TIME SKEW

With the conditioned transition probabilities derived in Section S1, there is a reflecting boundary at $m = 1$, which lets us set up a recurrence relation for the fixation-time moments. This derivation follows the method described by Keilson in Ref. [4]. Let $S_m(t)$ be the first passage time distribution from state m to state $m+1$. Clearly, $S_1(t)$ is an exponential distribution,

$$S_1(t) = \tilde{b}_1 e^{-\tilde{b}_1 t}. \tag{22}$$

From $m > 1$, the state $m+1$ can be reached either directly, with exponentially distributed times, or indirectly by first stepping backwards to $m-1$, returning to m, and then reaching $m+1$ at a latter time. Thus, the distributions $S_m(t)$ satisfy

$$S_m(t) = \tilde{b}_m e^{-(\tilde{b}_m+\tilde{d}_m)t} + \tilde{d}_m e^{-(\tilde{b}_m+\tilde{d}_m)t} * S_{m-1}(t) * S_m(t), \tag{23}$$

where the symbol $*$ denotes a convolution of distributions. This equation can be solved by Fourier transform to obtain

$$S_m(\omega) = \frac{\tilde{b}_m}{i\omega + \tilde{b}_m + \tilde{d}_m - \tilde{d}_m S_{m-1}(\omega)}. \tag{24}$$

We can compute a recurrence relation for the moments of the first passage time distributions $S_m(t)$ by differentiating Eq. 24. Let μ_m, ν_m and γ_m to be the first, second, and third moments of $S_m(t)$ respectively. Using the relations $\mu_m = iS'(\omega = 0)$, $\nu_m = iS''(\omega = 0)$, and $\gamma_m = iS'''(\omega = 0)$, we find that

$$\begin{aligned}
\mu_m &= \tilde{b}_m^{-1}(1 + \tilde{d}_m\mu_{m-1}),\\
\nu_m &= \tilde{b}_m^{-2}[\tilde{b}_m\tilde{d}_m\beta_{m-1} + 2(1 + \tilde{d}_m\mu_{m-1})^2],\\
\gamma_m &= \tilde{b}_m^{-3}[\tilde{b}_m^2\tilde{d}_m\gamma_{m-1} + 6\tilde{b}_m\tilde{d}_m\nu_{m-1}(1 + \tilde{d}_m\mu_{m-1}) + 6(1 + \tilde{d}_m\mu_{m-1})^3],
\end{aligned} \tag{25}$$

with boundary conditions $\mu_0 = \nu_0 = \gamma_0 = 0$. The recurrence relations in Eq. 25 give the moments of incremental first passage time distributions $S_m(t)$. The total fixation time, T is the sum of these incremental first passage times. Thus, the cumulants of T are the sum of the cumulants of the incremental times and the skew of T can be expressed as,

$$\kappa_3(N) = \left(\sum_{m=1}^{N-1} \gamma_m - 3\nu_m\mu_m + 2\mu_m^3\right) \Bigg/ \left(\sum_{m=1}^{N-1} \nu_m - \mu_m^2\right)^{3/2}. \tag{26}$$

Numerical computation of for $\kappa_3(N)$ requires calculating the $3N$ moments and carrying out the two sums in Eq. 26. By bottom-up tabulation of the incremental moments, this procedure can be completed in $\mathcal{O}(N)$ time, asymptotically faster than the eigenvalue decomposition and the exact series solution from visit statistics.

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

S4: ASYMPTOTIC ANALYSIS FOR THE 1D LATTICE

Neutral Fitness

As in the main text, we begin with the neutral fitness Moran process on a 1D lattice with periodic boundary conditions. In this case, the eigenvalues of the transition matrix describing the system are,

$$\lambda_m = \frac{1}{N} - \frac{1}{N}\cos\left(\frac{m\pi}{N}\right), \quad m = 1, 2, \ldots, N-1. \tag{27}$$

From the eigen-decomposition of the Markov birth-death process described in the main text, the standardized fixation-time cumulants are given by

$$\kappa_n(N) = (n-1)!\left(\sum_{m=1}^{N-1}\frac{1}{\lambda_m^n}\right)\Bigg/\left(\sum_{m=1}^{N-1}\frac{1}{\lambda_m^2}\right)^{n/2}. \tag{28}$$

Note that the constant factor $1/N$ cancels in Eq. 28, so we may equivalently consider rescaled eigenvalues $\lambda_m = 1 - \cos(m\pi/N)$. To derive the asymptotic cumulants, we compute the leading asymptotic behavior of sums

$$S_n = \sum_{m=1}^{N-1}\frac{1}{[1-\cos(m\pi/N)]^n}. \tag{29}$$

The function $(1-\cos x)^{-n}$ can be expanded as a Laurent series $\sum_{k=0}^{\infty} c_k(n)x^{2(k-n)}$, which is absolutely convergent for $x \neq 0$ in the interval $(-2\pi, 2\pi)$. So the sum S_n can then be expressed as

$$\begin{aligned} S_n &= \sum_{m=1}^{N-1}\sum_{k=0}^{\infty} c_k(n)\left(\frac{\pi m}{N}\right)^{2(k-n)} \\ &= \sum_{k=0}^{\infty} c_k(n)(N/\pi)^{2(n-k)} H_{N-1,2(n-k)} \\ &= \frac{c_0(n)\zeta(2n)}{\pi^{2n}}N^{2n} + \mathcal{O}(N^{2(n-1)}) \end{aligned} \tag{30}$$

where $H_{N,q} = \sum_{m=1}^{N} m^{-q}$ is the generalized harmonic number and in the last line we used the asymptotic approximation

$$H_{N,2q} = \begin{cases} \zeta(2q) + \mathcal{O}(N^{1-2q}) & q > 0, \\ \dfrac{N^{1-2q}}{2q+1} + \mathcal{O}(N^{-2q}) & q \leq 0. \end{cases} \tag{31}$$

It is easy to check that $c_0(n) = 2^n$. Now the cumulants are $\kappa_n(N) = (n-1)!S_n/S_2^{n/2}$, which for $N \to \infty$ are

$$\begin{aligned} \kappa_n &= (n-1)!\left(\frac{2^n\zeta(2n)}{\pi^{2n}}\right)\Bigg/\left(\frac{2^2\zeta(4)}{\pi^4}\right)^{n/2} \\ &= (n-1)!\,\frac{\zeta(2n)}{\zeta(4)^{n/2}}, \end{aligned} \tag{32}$$

as reported in the main text.

Non-neutral fitness

For non-neutral fitness, we showed in the main text that in the random walk approximation the fixation-time distribution is asymptotically normal. Here we use the visit statistics approach to prove this holds even with the variation in time spent in each state during the Moran process.

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

To prove the fixation-time distribution is normal, we derive bounds on the sums

$$S_n = \sum_{i_1,i_2,\ldots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})} \tag{33}$$

that appear in Eq. 15 to show that $\kappa_n(N) \to 0$ as $N \to \infty$. First note that $w^n_{i_1 i_2 \cdots i_n}(r,N) = (n-1)!\,V^n_{ii} \geq 1$ if $i_1 = i_2 = \cdots = i_n \equiv i$ so we can trivially bound S_n from below by the sum over unweighted diagonal elements. Furthermore, we claim that $w^n_{i_1 i_2 \cdots i_n}(r,N) \geq 0$ for all $i_1, i_2, \ldots, i_n$, since otherwise one could construct a birth-death process with negative fixation-time cumulants. But from the eigen-decomposition of the birth-death process, described in the main text, the fixation-time cumulants are all positive. Since $b_i + d_i$ is also positive for all i, the sums are bounded from above by the maximum value of $(b_i + d_i)^{-n}$ times the sum over the weighting factors. Putting these together, we obtain

$$\sum_{i=1}^{N-1} \frac{1}{(b_i+d_i)^n} \leq S_n \leq \left(\max_{1\leq i<N} \frac{1}{b_i+d_i}\right)^n \times \sum_{i_1,i_2,\ldots,i_n=1}^{N-1} w^n_{i_1 i_2 \cdots i_n}(r,N). \tag{34}$$

The Moran process on the 1D lattice has transition probabilities $b_i + d_i = (1+r)/(rm+N-m)$. Then, as $N \to \infty$, the lower bound is

$$\sum_{i=1}^{N-1} \frac{1}{(b_i+d_i)^n} = \frac{1}{(r+1)^n} \sum_{m=1}^{N-1} (rm+N-m)^n = \frac{1+r+r^2+\cdots+r^n}{(n+1)(1+r)^n} N^{n+1} + \mathcal{O}(N^n). \tag{35}$$

For the upper bound, first note that

$$\left(\max_{1\leq i<N} \frac{1}{b_i+d_i}\right)^n = [r(N-1)+1]^n = r^n N^n + \mathcal{O}(N^{n-1}). \tag{36}$$

The sums over the weighting factors give the (non-standardized) fixation-time cumulants corresponding to a process with $b_i + d_i = 1$ and uniform bias r. This is exactly the biased random walk model used to approximate the Moran process in the main text. It follows that

$$\sum_{i_1,i_2,\ldots,i_n=1}^{N-1} w^n_{i_1 i_2 \cdots i_n}(r,N) = (n-1)! \sum_{i=1}^{N-1} \left(\frac{1}{1-2\sqrt{r}/(r+1)\cos(m\pi/N)}\right)^n, \tag{37}$$

where the denominators in the second sum are the eigenvalues of the transition matrix for the biased random walk, $\lambda_m = 1 - 2\sqrt{r}/(r+1)\cos(m\pi/N)$. As in the main text, we can estimate the leading asymptotics of this sum by converting to an integral,

$$\sum_{i_1,i_2,\ldots,i_n=1}^{N-1} w^n_{i_1 i_2 \cdots i_n}(r,N) = \frac{N}{\pi} \int_0^\pi \frac{1}{(1-2\sqrt{r}/(1+r)\cos x)^n}\,\mathrm{d}x + \mathcal{O}(1). \tag{38}$$

Combining the results from Eqs. 34–36 and 38 we arrive at

$$\frac{1+r+r^2+\cdots+r^n}{(n+1)(1+r)^n} N^{n+1} + \mathcal{O}(N^n) \leq S_n \leq \frac{N^{n+1}}{\pi} \int_0^\pi \frac{1}{(1-2\sqrt{r}/(1+r)\cos x)^n}\,\mathrm{d}x + \mathcal{O}(N^n). \tag{39}$$

For each n, our upper and lower bounds have the same asymptotic scaling as a power of N, with different r-dependent coefficients. Using these results together in Eq. 15, it follows that for $N \gg 1$, the cumulants to leading order are

$$\kappa_n(N) = C_n(r) \frac{1}{N^{(n-2)/2}} + \mathcal{O}(N^{-n/2}), \tag{40}$$

where $C_n(r)$ is a fitness-dependent constant. Thus, indeed $\kappa_n(N) \to 0$ as $N \to \infty$.

This result confirms the claim made in the main text. Even with heterogeneity in the time spent in each state, the skew and higher-order cumulants of the fixation time vanish asymptotically. Therefore, the Moran Birth-death process on the 1D lattice with non-neutral fitness $r > 1$ has an asymptotically normal fixation-time distribution. The normal distribution is universal, independent of fitness level for this population structure.

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

S5: ASYMPTOTIC ANALYSIS FOR THE COMPLETE GRAPH

Non-neutral fitness

In the main text we predicted that the asymptotic fixation-time distribution for the Moran Birth-death process on the complete graph is a convolution of two Gumbel distributions by applying our intuition from coupon collection. Furthermore, our calculation of the fixation-time cumulants to arbitrary order in the large (but finite) fitness limit agrees with this prediction. Surprisingly, numerical calculations using the recurrence relation formulated above and direct simulations of the Moran process indicate that this result holds for all $r > 1$. In this section we prove, using the visit statistics formulation, that the asymptotic skew of the fixation time for $r > 1$ is identical to that of a convolution of Gumbel distributions. Based on our numerical evidence, we conjecture that an analogous calculation holds to all orders. The below calculation also shows why the coupon collection heuristic works: the asymptotically dominant terms come exclusively from the regions near fixation ($m = N - 1$) and near the beginning of the process when a single mutant is introduced into the system ($m = 1$).

As for the 1D lattice, we want to derive the asymptotic behavior of the sums

$$S_n = \sum_{i_1,i_2,\ldots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})}, \tag{41}$$

where the transition probabilities b_i and d_i are those for the Moran process on the complete graph,

$$b_i + d_i = \frac{(1+r)i(N-i)}{(N-1)(ri+N-i)}. \tag{42}$$

To start, consider the sums Eq. 41, but with two indices i_1 and i_2 constrained to integers from αN to $(1-\alpha)N$ for $1/2 > \alpha > 0$. This sum may be written as

$$S^\alpha_n = \sum_{i_1,i_2=\alpha N}^{(1-\alpha)N} \sum_{i_3,\cdots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})}. \tag{43}$$

Now we may apply the upper bound in Eq. 34, except for the sums restricted to $\alpha N < i, j < (1-\alpha)N$ the maximum of $(b_i + d_i)^{-1}$ can be restricted to this range,

$$\begin{aligned} S^\alpha_n &\le \left(\max_{1<i<N} \frac{1}{b_i+d_i}\right)^{n-2} \times \left(\max_{\alpha N<i<(1-\alpha)N} \frac{1}{b_i+d_i}\right)^2 \times \sum_{i_1,i_2,\cdots,i_n=1}^{N} w^n_{i_1 i_2 \cdots i_n}(r,N) \\ &= N^{n-1}\left\{\left(\frac{r}{1+r}\right)^{n-2}\left(\frac{r(1-\alpha)+\alpha}{(1+r)(1-\alpha)\alpha}\right)^2 \times \frac{1}{\pi}\int_0^\pi \frac{1}{(1-2\sqrt{r}/(1+r)\cos x)^n}\,\mathrm{d}x\right\} + \mathcal{O}(N^{n-2}). \end{aligned} \tag{44}$$

In the second line we used the integral approximation from Eq. 38 and evaluated the maximum of $(b_i + d_i)^{-1}$ over the indicated intervals. Since we constructed $w^n_{i_1 i_2 \cdots i_n}(r,N)$ to be symmetric, this upper bound holds for any permutation of the indices in Eq. 43.

We now consider the same sums but with $1 < i_1 < \alpha N$ or $(1-\alpha)N < i_1 < N-1$,

$$S^{\alpha,1}_n = \sum_{i_1=1}^{\alpha N} \sum_{i_2=\alpha N}^{(1-\alpha)N} \sum_{i_3,\cdots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})}. \tag{45}$$

and

$$S^{\alpha,2}_n = \sum_{i_1=(1-\alpha)N}^{N-1} \sum_{i_2=\alpha N}^{(1-\alpha)N} \sum_{i_3,\cdots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})}. \tag{46}$$

These sums can be estimated using the same upper bound, but without extending the sum on $w^n_{i_1 i_2 \cdots i_n}(r,N)$ to the entire domain. Specifically,

$$S^{\alpha,1}_n \le N^{n-1}\left\{\left(\frac{r}{1+r}\right)^{n-1}\left(\frac{r(1-\alpha)+\alpha}{(1+r)(1-\alpha)\alpha}\right)\right\} \times \sum_{i_1=1}^{\alpha N} \sum_{i_2=\alpha N}^{(1-\alpha)N} \sum_{i_3,\cdots,i_n=1}^{N-1} w^n_{i_1 i_2 \cdots i_n}(r,N) + \mathcal{O}(N^{n-2}). \tag{47}$$

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Note that the weighting factors fall off exponentially away from the diagonal elements. This is because the visit numbers in the biased random walk become only very weakly correlated if the states are far away. Thus, the sum in Eq. 47 over terms away from the diagonal elements converges to a constant as $N \to \infty$. We have verified this explicitly for $w^2_{ij}(r,N)$ and $w^3_{ijk}(r,n)$. The series $S^{\alpha,2}_n$ is similarly bounded, as are all sums of the form Eq. 45 or 46 with the indices permuted.

The remaining terms in S_n involve all indices in either $[1,\alpha N]$ or $[(1-\alpha)N, N-1]$. If not all indices are in the same interval, the weighting factors are exponentially small: the visit numbers near $m=1$ are uncorrelated with those near $m=N-1$. Thus each term in the sum is exponentially suppressed and doesn't contribute to S_n asymptotically. With this observation only two parts of the sum remain: those with bounds $1 \le i_1, i_2 \dots i_n \le \alpha N$ or $(1-\alpha)N \le i_1, i_2 \dots i_n \le N-1$. We call the sums with these bounds S^{c1}_n and S^{c2}_n respectively. As we will see below, the sums over these regions have leading order $\mathcal{O}(N^n)$. Since all the above terms are order $\mathcal{O}(N^{n-1})$ or smaller, the asymptotic behavior of the cumulants is entirely determined by these regions near the beginning and end of the process, i.e. the coupon collection regions. The fact that we can restrict the sums to this region allows us to make approximations that do not change the asymptotic behavior, but make the sums easier to carry out. For instance, in S^{c1}_2, we can set $r^N - r^i \to r^N$ and $(N-i) \to N$, since the indices run only up to αN. This gives

$$\begin{aligned} S^{c1}_2 &= \frac{N^2}{(r-1)^2}\left\{\sum_{i=1}^{\alpha N}\frac{(r^i-1)^2}{i^2 r^{2i}} + 2\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1}\frac{(r^j-1)^2}{ij\, r^{i+j}}\right\} + \mathcal{O}(N) \\ &= \frac{N^2\zeta(2)}{(r-1)^2} + \mathcal{O}(N), \end{aligned} \tag{48}$$

for $N \gg 1$. A similar approximation shows $S^{c2}_2 = r^2 N^2 \zeta(2)/(r-1)^2$. For the third order sums, we find

$$\begin{aligned} S^{c1}_3 = 2\frac{N^3}{(r-1)^3}\Bigg\{&\sum_{i=1}^{\alpha N}\frac{(r^i-1)^3}{i^3 r^{3i}} + 3\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1}\frac{(r^j-1)^3}{ij^2 r^{i+2j}} \\ &+ 3\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1}\frac{(r^i-1)(r^j-1)^2}{i^2 j r^{2i+j}} + +6\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}\frac{(r^j-1)(r^k-1)^2}{ijk\, r^{i+j+k}}\Bigg\} + \mathcal{O}(N^2) \\ = 2\frac{N^3\zeta(3)}{(r-1)^3} + \mathcal{O}(N^2), \end{aligned} \tag{49}$$

for $N \gg 1$. Again the other sum, with indices near $N-1$, is identical up to a factor of r^3, $S^{c2}_3 = r^3\zeta(3)/(r-1)^3$. Overall, we have that

$$S_2 = \frac{N^2(1+r^2)\zeta(2)}{(r-1)^2} + \mathcal{O}(N) \qquad \text{and} \qquad S_3 = \frac{2N^3(1+r^3)\zeta(3)}{(r-1)^3} + \mathcal{O}(N^2). \tag{50}$$

Then the asymptotic skew is given by

$$\kappa_3 = \frac{2(1+r^3)\zeta(3)}{(r-1)^3}\Bigg/\left(\frac{(1+r^2)\zeta(2)}{(r-1)^2}\right)^{3/2} = \frac{1+r^3}{(1+r^2)^{3/2}} \times \frac{2\zeta(3)}{\zeta(2)^{3/2}}, \tag{51}$$

which is exactly the skew corresponding to the convolution of Gumbel distributions with relative weighting given by the fitness, $G + rG$. While evaluating the series to higher orders is increasingly difficult, our simulations and the large-fitness approximation suggest this result holds to all orders and that indeed, the asymptotic fixation-time distribution is a weighted convolution of Gumbel distributions.

Neutral fitness with truncation

As discussed in the main text, the neutral fitness Moran process on the complete graph has a fixation-time skew that depends on the level of truncation. That is, the time T_α it takes for the process to reach αN mutants, where $0 \le \alpha \le 1$, has a distribution whose skew depends on α. Here we show that the $\alpha \to 0$ limit of the fixation-time skew equals $\sqrt{3}$.

To start, we take the neutral fitness limit of the weighting factors to obtain

$$w^2_{ij}(1,\alpha N) = \frac{4k^2(\alpha N - j)^2}{\alpha^2 N^2} \tag{52}$$

bioRxiv preprint doi: https://doi.org/10.1101/496380; this version posted December 13, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

for $i \geq j$ and

$$w^3_{ijk}(1, \alpha N) = \frac{16kl^2(\alpha N - j)^2(\alpha N - k)}{\alpha^3 N^3} \tag{53}$$

for $i \geq j \geq k$, again with the expressions for other orderings obtained by permuting the indices accordingly. The neutral fitness Moran process on the complete graph has transition probabilities $b_i + d_i = 2(Ni - i^2)/(N^2 - N)$. Since we are computing the truncated fixation-time skew, we use Eq. 15, but cut the sums off at αN. In this case, these sums are dominated by the off-diagonal terms, so that

$$\begin{aligned} S_2 = \sum_{i,j=1}^{\alpha N} \frac{w^2_{ij}(1, \alpha N)}{(b_i + d_i)(b_j + d_j)} &= 2\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1} \frac{j(\alpha N - i)^2(N-1)^2}{\alpha^2 i(N-i)(N-j)} + \mathcal{O}(N^3) \\ &= 2\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1} \frac{j(\alpha N - i)^2}{\alpha^2 i} + \mathcal{O}(N^3) \\ &= \frac{\alpha^2 N^4}{12} + \mathcal{O}(N^3), \end{aligned} \tag{54}$$

where in the second line we approximated $N - i$ and $N - j$ by N. This approximation is exact in the limit $\alpha \to 0$ since the upper limit on the sum is αN, which is much smaller than N. Using analogous approximations, we find

$$\begin{aligned} S_3 = \sum_{i,j,k=1}^{\alpha N} \frac{w^2_{ijk}(1, \alpha N)}{(b_i + d_i)(b_j + d_j)(b_k + d_k)} &= 6\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} \frac{2k(\alpha N - i)^2(\alpha N - j)(N-1)^3}{\alpha^3 i(N-i)(N-j)(N-k)} + \mathcal{O}(N^5) \\ &= 12\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} \frac{k(\alpha N - i)^2(\alpha N - j)}{\alpha^2 i} + \mathcal{O}(N^5) \\ &= \frac{\alpha^3 N^6}{24} + \mathcal{O}(N^5). \end{aligned} \tag{55}$$

The asymptotic fixation-time skew as $\alpha \to 0$ is therefore

$$\kappa_3 = \frac{\alpha^2 N^6/24}{(\alpha^2 N^4/12)^{3/2}} = \sqrt{3}, \tag{56}$$

as claimed in the main text. This value agrees perfectly with our numerical calculations, which show the above approximation breaks down when $\alpha \approx 1/2$. Above this threshold, the random walk causes mixing between the two coupon collection regions, thereby lowering the overall skew of the fixation-time distribution toward the $\alpha = 1$ value of $\kappa_3 = 6\sqrt{3}(10 - \pi^2)/(\pi^2 - 9)^{3/2} \approx 1.6711$.

[1] Kaveh K, Komarova NL, Kohandel M (2015) The duality of spatial death–birth and birth–death processes and limitations of the isothermal theorem. *Royal Society open science* 2(4):140465.

[2] da Fonseca C, Petronilho J (2001) Explicit inverses of some tridiagonal matrices. *Linear Algebra and its Applications* 325(1):7 – 21.

[3] Kemeny JG, Snell JL (1983) *Finite Markov chains: with a new appendix "Generalization of a fundamental matrix"*. (Springer).

[4] Keilson J (1965) A review of transient behavior in regular diffusion and birth-death processes. part ii. *Journal of Applied Probability* 2(2):405–428.