Supplemental Material for "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: May 31, 2019)

This Supplemental Material is devoted to providing rigorous mathematical derivations supporting the results obtained 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 for the visit statistics approach (formulated in the main text Appendix, 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 S2). In Sections S3 and S4 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. Finally, in Section S5 we give further details regarding the extension of our results to the two-fitness Moran process.

S1. BIRTH-DEATH MARKOV CHAIN CONDITIONED ON FIXATION

For both the numerical recurrence relation and the visit statistics approach, it is useful to consider the birth-death Markov chain conditioned on hitting N, which has an identical fixation-time distribution to the unconditioned process. This Markov chain has new conditioned transition probabilities denoted $\tilde{b}_m$ 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

$$\begin{aligned}\tilde{b}_m &= \frac{\mathcal{P}(X_{t+1} = m+1 \text{ AND } X_t = m \text{ AND } X_\infty = N)}{\mathcal{P}(X_t = m \text{ AND } X_\infty = N)} \\ &= \frac{\mathcal{P}(X_\infty = N | X_t = m+1)}{\mathcal{P}(X_\infty = N | X_t = m)} \mathcal{P}(X_{t+1} = m+1 | X_t = m) \\ &= \frac{\mathcal{P}(X_\infty = N | X_t = m+1)}{\mathcal{P}(X_\infty = N | X_t = m)} b_m,\end{aligned} \tag{S1}$$

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

$$\tilde{d}_m = \frac{\mathcal{P}(X_\infty = N | X_t = m-1)}{\mathcal{P}(X_\infty = N | X_t = m)} d_m. \tag{S2}$$

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:

$$\begin{aligned}\tilde{b}_m + \tilde{d}_m &= \frac{\mathcal{P}(X_{t+1} = m+1 \text{ AND } X_t = m \text{ AND } X_\infty = N) + \mathcal{P}(X_{t+1} = m-1 \text{ AND } X_t = m \text{ AND } X_\infty = N)}{\mathcal{P}(X_t = m \text{ AND } X_\infty = N)} \\ &= 1 - \frac{\mathcal{P}(X_{t+1} = m \text{ AND } X_t = m \text{ AND } X_\infty = N)}{\mathcal{P}(X_t = m \text{ AND } X_\infty = N)} \\ &= 1 - \mathcal{P}(X_{t+1} = m | X_t = m) \\ &= b_m + d_m.\end{aligned} \tag{S3}$$

This invariance, along with Eqs. (S1) and (S2), shows that conditioning the Markov chain is equivalent to a similarity transformation on the transient transition matrix with a diagonal change of basis:

$$\tilde{\Omega}_{\rm tr} = S\,\Omega_{\rm tr}\,S^{-1} \qquad S_{mn} = \mathcal{P}(X_\infty = N|X_t = m)\delta_{m,n}, \tag{S4}$$

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

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

$$\mathcal{P}(X_\infty = N|X_t = m) = \frac{1-1/r^m}{1-1/r^N}, \tag{S5}$$

so that

$$\tilde{b}_m = \frac{r^{m+1}-1}{r^{m+1}-r} b_m, \qquad \tilde{d}_m = \frac{r^m - r}{r^m - 1} d_m. \tag{S6}$$

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 two-fitness Moran Birth-Death model discussed in the main text fixation probabilities derived by Kaveh et al. [1] can be used together with Eq. (S4) to condition the Markov chain on fixation.

S2. 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. [2]. Let $S_m(t)$ be the first passage time densities from state m to state $m+1$. Clearly, $S_1(t)$ has an exponential distribution,

$$S_1(t) = \tilde{b}_1 e^{-\tilde{b}_1 t}. \tag{S7}$$

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 densities $S_m(t)$ satisfy

$$S_m(t) = \tilde{b}_m e^{-(\tilde{b}_m+\tilde{d}_m)t} + \tilde{d}_m e^{-(\tilde{b}_m+\tilde{d}_m)t} * S_{m-1}(t) * S_m(t), \tag{S8}$$

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

$$S_m(\omega) = \frac{\tilde{b}_m}{\tilde{b}_m + \tilde{d}_m - \tilde{d}_m S_{m-1}(\omega) - i\omega}. \tag{S9}$$

We can compute a recurrence relation for the moments of the first passage time densities $S_m(t)$ by differentiating Eq. (S9). Let μ_m, ν_m and γ_m to be the first, second, and third moments of $S_m(t)$ respectively. Using the relations $\nu_m = -iS'(\omega = 0)$, $\xi_m = (-i)^2 S''(\omega = 0)$, and $\zeta_m = (-i)^3 S'''(\omega = 0)$, we find that

$$\begin{aligned} \nu_m &= \tilde{b}_m^{-1}(1 + \tilde{d}_m \nu_{m-1}), \\ \xi_m &= \tilde{b}_m^{-2}[\tilde{b}_m \tilde{d}_m \xi_{m-1} + 2(1 + \tilde{d}_m \nu_{m-1})^2], \\ \zeta_m &= \tilde{b}_m^{-3}[\tilde{b}_m^2 \tilde{d}_m \zeta_{m-1} + 6\tilde{b}_m \tilde{d}_m \xi_{m-1}(1 + \tilde{d}_m \nu_{m-1}) + 6(1 + \tilde{d}_m \nu_{m-1})^3], \end{aligned} \tag{S10}$$

with boundary conditions $\nu_0 = \xi_0 = \zeta_0 = 0$. The recurrence relations in Eq. (S10) 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,

$$\kappa_3(N) = \left(\sum_{m=1}^{N-1} \zeta_m - 3\xi_m \nu_m + 2\nu_m^3 \right) \Bigg/ \left(\sum_{m=1}^{N-1} \xi_m - \nu_m^2 \right)^{3/2}. \tag{S11}$$

Numerical computation of for $\kappa_3(N)$ requires calculating the $3N$ moments and carrying out the two sums in Eq. (S11). 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.

S3. ASYMPTOTIC ANALYSIS FOR THE 1D LATTICE

A. 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,

$$\lambda_m = \frac{2}{N} - \frac{2}{N}\cos\left(\frac{m\pi}{N}\right), \quad m = 1, 2, \ldots, N-1. \tag{S12}$$

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

$$\kappa_n(N) = (n-1)!\left(\sum_{m=1}^{N-1}\frac{1}{\lambda_m^n}\right) \Bigg/ \left(\sum_{m=1}^{N-1}\frac{1}{\lambda_m^2}\right)^{n/2}. \tag{S13}$$

Note that the constant factor $2/N$ cancels in Eq. (S13), 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

$$S_n = \sum_{m=1}^{N-1}\frac{1}{[1-\cos(m\pi/N)]^n}. \tag{S14}$$

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

$$\begin{aligned} S_n &= \sum_{m=1}^{N-1}\sum_{k=0}^{\infty} c_k(n)\left(\frac{\pi m}{N}\right)^{2(k-n)} \\ &= \sum_{k=0}^{\infty} c_k(n)(N/\pi)^{2(n-k)}H_{N-1,2(n-k)} \\ &= \frac{c_0(n)\zeta(2n)}{\pi^{2n}}N^{2n} + \mathcal{O}(N^{2(n-1)}) \end{aligned} \tag{S15}$$

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

$$H_{N,2q} = \begin{cases} \zeta(2q) + \mathcal{O}(N^{1-2q}) & q > 0, \\ \dfrac{N^{1-2q}}{2q+1} + \mathcal{O}(N^{-2q}) & q \leq 0. \end{cases} \tag{S16}$$

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

$$\begin{aligned} \kappa_n &= (n-1)!\left(\frac{2^n\zeta(2n)}{\pi^{2n}}\right) \Bigg/ \left(\frac{2^2\zeta(4)}{\pi^4}\right)^{n/2} \\ &= (n-1)!\,\frac{\zeta(2n)}{\zeta(4)^{n/2}}, \end{aligned} \tag{S17}$$

as reported in the main text.

B. 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 when the variation in time spent in each state is accounted for. From the visit statistics formulation, the standardized cumulants of the fixation time (starting from a single initial mutant) can be written as,

$$\kappa_n(N) = \left(\sum_{i_1,i_2,\ldots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})} \right) \Bigg/ \left(\sum_{i,j=1}^{N-1} \frac{w^2_{ij}(r,N)}{(b_i+d_i)(b_j+d_j)} \right)^{n/2}, \tag{S18}$$

where $w^n_{i_1 i_2 \cdots i_n}(r,N)$ are the weighting factors that depend on the visit statistics of a biased random walk. To prove the fixation-time distribution is normal, we derive bounds on the sums

$$S_n = \sum_{i_1,i_2,\ldots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})} \tag{S19}$$

that appear in Eq. (S18) and 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 $i_1 = i_2 = \cdots = i_n \equiv i$. Furthermore, we claim that $w^n_{i_1 i_2 \cdots i_n}(r,N) \geq 0$ for all $i_1, i_2, \ldots, i_n$. If this were not the case, one could construct a birth-death process with negative fixation-time cumulants by choosing $b_i + d_i$ appropriately. But we know the fixation-time cumulants are positive from the eigen-decomposition described in the main text. With these observations, we can bound S_n from below by the sum over unweighted diagonal elements. Similarly, 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

$$\sum_{i=1}^{N-1} \frac{1}{(b_i+d_i)^n} \leq S_n \leq \left(\max_{1\leq i<N} \frac{1}{b_i+d_i} \right)^n \times \sum_{i_1,i_2,\ldots,i_n=1}^{N-1} w^n_{i_1 i_2 \cdots i_n}(r,N). \tag{S20}$$

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

$$\sum_{i=1}^{N-1} \frac{1}{(b_i+d_i)^n} = \frac{1}{(r+1)^n} \sum_{m=1}^{N-1} (rm+N-m)^n = \frac{1+r+r^2+\cdots+r^n}{(n+1)(1+r)^n} N^{n+1} + \mathcal{O}(N^n). \tag{S21}$$

For the upper bound, first note that

$$\left(\max_{1\leq i<N} \frac{1}{b_i+d_i} \right)^n = [r(N-1)+1]^n = r^n N^n + \mathcal{O}(N^{n-1}). \tag{S22}$$

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 as $N \to \infty$,

$$\sum_{i_1,i_2,\ldots,i_n=1}^{N-1} w^n_{i_1 i_2 \cdots i_n}(r,N) = (n-1)! \sum_{i=1}^{N-1} \left(\frac{1}{1-2\sqrt{r}/(r+1)\cos(m\pi/N)} \right)^n, \tag{S23}$$

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,

$$\sum_{i_1,i_2,\ldots,i_n=1}^{N-1} w^n_{i_1 i_2 \cdots i_n}(r,N) = \frac{N}{\pi} \int_0^\pi \frac{(n-1)!}{(1-2\sqrt{r}/(1+r)\cos x)^n}\,\mathrm{d}x + \mathcal{O}(1). \tag{S24}$$

Combining the results from Eqs. (S20)–(S22) and (S24) we arrive at

$$\frac{1+r+r^2+\cdots+r^n}{(n+1)(1+r)^n} N^{n+1} + \mathcal{O}(N^n) \leq S_n \leq \frac{N^{n+1}}{\pi} \int_0^\pi \frac{(n-1)!}{(1-2\sqrt{r}/(1+r)\cos x)^n}\,\mathrm{d}x + \mathcal{O}(N^n). \tag{S25}$$

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. (S18), it follows that for $N \gg 1$, the cumulants to leading order are

$$\kappa_n(N) = C_n(r) \frac{1}{N^{(n-2)/2}} + \mathcal{O}(N^{-n/2}), \tag{S26}$$

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.

S4. ASYMPTOTIC ANALYSIS FOR THE COMPLETE GRAPH

A. 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 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

$$S_n = \sum_{i_1,i_2,\ldots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})}, \tag{S27}$$

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

$$b_i + d_i = \frac{(1+r)i(N-i)}{(N-1)(ri+N-i)} \tag{S28}$$

and the weights $w^2_{ij}(r,N)$ and $w^3_{ijk}(r,N)$ are respectively given by

$$w^2_{ij}(r,N) = \frac{(r+1)^2(r^j-1)^2(r^N-r^i)^2}{r^{i+j}\,(r-1)^2(r^N-1)^2} \tag{S29}$$

for $i > j$ and

$$w^3_{ijk}(r,N) = 2\frac{(r+1)^3(r^k-1)^2(r^j-1)(r^N-r^i)^2(r^N-r^j)}{r^{i+j+k}\,(r-1)^3(r^N-1)^3}, \tag{S30}$$

for $i > j > k$. The expressions for different orderings of indices are the same but with the indices permute appropriately so that w^2_{ij} and w^3_{ijk} are perfectly symmetric.

To start, consider the sums Eq. (S27), 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

$$S^\alpha_n = \sum_{i_1,i_2=\alpha N}^{(1-\alpha)N} \sum_{i_3,\cdots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})}. \tag{S31}$$

Now we may apply the upper bound in Eq. (S20), but for the sums restricted to $\alpha N < i,j < (1-\alpha)N$, the maximum of $(b_i+d_i)^{-1}$ can also be restricted to this range,

$$\begin{aligned} S^\alpha_n &\le \left(\max_{1<i<N} \frac{1}{b_i+d_i}\right)^{n-2} \times \left(\max_{\alpha N<i<(1-\alpha)N} \frac{1}{b_i+d_i}\right)^2 \times \sum_{i_1,i_2,\cdots,i_n=1}^{N} w^n_{i_1 i_2 \cdots i_n}(r,N) \\ &= N^{n-1}\left\{\left(\frac{r}{1+r}\right)^{n-2}\left(\frac{r(1-\alpha)+\alpha}{(1+r)(1-\alpha)\alpha}\right)^2 \times \frac{1}{\pi}\int_0^\pi \frac{(n-1)!}{(1-2\sqrt{r}/(1+r)\cos x)^n}\,\mathrm{d}x\right\} + \mathcal{O}(N^{n-2}). \end{aligned} \tag{S32}$$

In the second line we used the integral approximation from Eq. (S24) 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. (S31).

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

$$S_n^{\alpha,1} = \sum_{i_1=1}^{\alpha N} \sum_{i_2=\alpha N}^{(1-\alpha)N} \sum_{i_3,\cdots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})}. \tag{S33}$$

and

$$S_n^{\alpha,2} = \sum_{i_1=(1-\alpha)N}^{N-1} \sum_{i_2=\alpha N}^{(1-\alpha)N} \sum_{i_3,\cdots,i_n=1}^{N-1} \frac{w^n_{i_1 i_2 \cdots i_n}(r,N)}{(b_{i_1}+d_{i_1})(b_{i_2}+d_{i_2})\cdots(b_{i_n}+d_{i_n})}. \tag{S34}$$

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,

$$S_n^{\alpha,1} \le N^{n-1} \left\{ \left(\frac{r}{1+r}\right)^{n-1} \left(\frac{r(1-\alpha)+\alpha}{(1+r)(1-\alpha)\alpha}\right) \right\} \times \sum_{i_1=1}^{\alpha N} \sum_{i_2=\alpha N}^{(1-\alpha)N} \sum_{i_3,\cdots,i_n=1}^{N-1} w^n_{i_1 i_2 \cdots i_n}(r,N) + \mathcal{O}(N^{n-2}). \tag{S35}$$

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 from each other. Thus, the sum in Eq. (S35) 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 $S_n^{\alpha,2}$ is similarly bounded, as are all sums of the form Eq. (S33) or (S34) 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 $1 \le i_1, i_2 \ldots i_n \le \alpha N$ or $(1-\alpha)N \le i_1, i_2 \ldots i_n \le N-1$. We call the sums with these bounds S_n^{c1} and S_n^{c2} 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 leading asymptotics, but make the sums easier to carry out. For instance, in S_2^{c1}, 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

$$\begin{aligned} S_2^{c1} &= \frac{N^2}{(r-1)^2} \left\{ \sum_{i=1}^{\alpha N} \frac{(r^i-1)^2}{i^2 r^{2i}} + 2 \sum_{i=1}^{\alpha N} \sum_{j=1}^{i-1} \frac{(r^j-1)^2}{ij\, r^{i+j}} \right\} + \mathcal{O}(N) \\ &= \frac{N^2 \zeta(2)}{(r-1)^2} + \mathcal{O}(N), \end{aligned} \tag{S36}$$

for $N \gg 1$. A similar calculation shows $S_2^{c2} = r^2 N^2 \zeta(2)/(r-1)^2$. For the third order sums, we find

$$\begin{aligned} S_3^{c1} &= 2\frac{N^3}{(r-1)^3} \Bigg\{ \sum_{i=1}^{\alpha N} \frac{(r^i-1)^3}{i^3 r^{3i}} + 3 \sum_{i=1}^{\alpha N} \sum_{j=1}^{i-1} \frac{(r^j-1)^3}{ij^2 r^{i+2j}} \\ &\qquad + 3 \sum_{i=1}^{\alpha N} \sum_{j=1}^{i-1} \frac{(r^i-1)(r^j-1)^2}{i^2 j r^{2i+j}} + 6 \sum_{i=1}^{\alpha N} \sum_{j=1}^{i-1} \sum_{k=1}^{j-1} \frac{(r^j-1)(r^k-1)^2}{ijk\, r^{i+j+k}} \Bigg\} + \mathcal{O}(N^2) \\ &= 2\frac{N^3 \zeta(3)}{(r-1)^3} + \mathcal{O}(N^2), \end{aligned} \tag{S37}$$

for $N \gg 1$. Again the other sum, with indices near $N-1$, is identical up to a factor of r^3, $S_3^{c2} = 2r^3 N^3 \zeta(3)/(r-1)^3$. Overall, we have that

$$S_2 = \frac{N^2(1+r^2)\zeta(2)}{(r-1)^2} + \mathcal{O}(N) \qquad \text{and} \qquad S_3 = \frac{2N^3(1+r^3)\zeta(3)}{(r-1)^3} + \mathcal{O}(N^2). \tag{S38}$$

The asymptotic skew is given by

$$\kappa_3 = \frac{2(1+r^3)\zeta(3)}{(r-1)^3} \bigg/ \left(\frac{(1+r^2)\zeta(2)}{(r-1)^2}\right)^{3/2} = \frac{1+r^3}{(1+r^2)^{3/2}} \times \frac{2\zeta(3)}{\zeta(2)^{3/2}}, \tag{S39}$$

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.

B. 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

$$w_{ij}^2(1, \alpha N) = \frac{4j^2(\alpha N - i)^2}{\alpha^2 N^2} \tag{S40}$$

for $i \ge j$ and

$$w_{ijk}^3(1, \alpha N) = \frac{16jk^2(\alpha N - i)^2(\alpha N - j)}{\alpha^3 N^3} \tag{S41}$$

for $i \ge j \ge k$, 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. (S18), but cut the sums off at αN. In this case, these sums are dominated by the off-diagonal terms, so that

$$\begin{aligned} S_2 = \sum_{i,j=1}^{\alpha N} \frac{w_{ij}^2(1,\alpha N)}{(b_i+d_i)(b_j+d_j)} &= 2\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1} \frac{j(\alpha N - i)^2(N-1)^2}{\alpha^2 i(N-i)(N-j)} + \mathcal{O}(N^3) \\ &= 2\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1} \frac{j(\alpha N - i)^2}{\alpha^2 i} + \mathcal{O}(N^3) \\ &= \frac{\alpha^2 N^4}{12} + \mathcal{O}(N^3), \end{aligned} \tag{S42}$$

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, αN, is much smaller than N. Using analogous approximations, we find

$$\begin{aligned} S_3 = \sum_{i,j,k=1}^{\alpha N} \frac{w_{ijk}^2(1,\alpha N)}{(b_i+d_i)(b_j+d_j)(b_k+d_k)} &= 6\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} \frac{2k(\alpha N - i)^2(\alpha N - j)(N-1)^3}{\alpha^3 i(N-i)(N-j)(N-k)} + \mathcal{O}(N^5) \\ &= 12\sum_{i=1}^{\alpha N}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} \frac{k(\alpha N - i)^2(\alpha N - j)}{\alpha^2 i} + \mathcal{O}(N^5) \\ &= \frac{\alpha^3 N^6}{24} + \mathcal{O}(N^5). \end{aligned} \tag{S43}$$

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

$$\kappa_3 = \frac{\alpha^3 N^6/24}{(\alpha^2 N^4/12)^{3/2}} = \sqrt{3}, \tag{S44}$$

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 $\kappa_3 = 6\sqrt{3}(10 - \pi^2)/(\pi^2 - 9)^{3/2} \approx 1.6711$.

S5. FIXATION-TIME DISTRIBUTIONS IN THE TWO-FITNESS MORAN PROCESS

As discussed in the main text, the two-fitness Birth-Death (BD) Moran process has the same family of fixation-time distributions as the Birth-death (Bd) process with only one fitness level. Here we provide further details leading to this conclusion. In particular, we give the transition probabilities for the two-fitness model and describe how the calculations from the main text generalize to this system. Here r is the fitness level during the birth step, while $\tilde{r}$ is the fitness level during the death step in the Moran process.

A. One-dimensional lattice

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

$$b_m = \frac{r}{rm + N - m}\frac{\tilde{r}}{1+\tilde{r}}, \qquad d_m = \frac{1}{r\tilde{r}} b_m, \tag{S45}$$

for $1 < m < N-1$. The probabilities are different when there is only one mutant or non-mutant ($m = 1$ or $m = N-1$ respectively). In these cases the nodes 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. (S45).

The two-fitness Moran BD model on the 1D lattice differs from the previously considered 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 expressions for the cumulants, Eqs. (S13) and (S18). 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.

B. Complete Graph

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

$$b_m = \frac{rm}{rm + N - m} \cdot \frac{\tilde{r}(N-m)}{\tilde{r}(N-m) + m - 1}, \qquad d_m = \frac{N-m}{rm + N - m} \cdot \frac{m}{\tilde{r}(N-m-1) + m}. \tag{S46}$$

In this case, the ratio of transition probabilities is m-dependent, but $b_m/d_m \to r\tilde{r}$ as $N \to \infty$, 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. (S13) with $\lambda_m \to b_m + d_m$. As $N \to \infty$, the cumulants become

$$\kappa_n = \frac{1 + r^n/\tilde{r}^n}{(1 + r^2/\tilde{r}^2)^{n/2}} \cdot \frac{(n-1)!\zeta(n)}{\zeta(2)^{n/2}}, \tag{S47}$$

identical to the Moran Bd process on the complete graph, with $r \to r/\tilde{r}$. Numerical calculations (using the moment recurrence relation derived in Section S2) again indicate this expression for the cumulants holds for all r, not just in the large 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 numerics indicate there is an entire family of distributions dependent on $r = 1/\tilde{r}$.

[1] K. Kaveh, N. L. Komarova, and M. Kohandel, Royal Society Open Science **2**, 140465 (2015).
[2] J. Keilson, Journal of Applied Probability **2**, 405 (1965).