

## Supplementary Notes

Here we sketch the derivations of eq (1) for circulations and eq (2) for superstars. We give a brief discussion of complexity results for frequency-dependent selection and the computation underlying our results for directed cycles. We close with a discussion of our assumptions about mutation rate and the interpretations of fitness which these results can accommodate.

### Evolution on graphs is a Markov process.

Let  $G$  be a graph whose adjacency matrix is given by  $W$ . Let  $\mathbf{P} \subset \mathbf{V}$  be the set of vertices occupied by a mutant at some iteration.  $\mathbf{P}$  represents a state of the typical Markov chain  $E_G$  which arises on an evolutionary graph. Analogously, the states  $P = \{1, 2, \dots, N\}$  are the typical states of the Moran process  $M$ .

(For two types of individuals, the states of the explicit Markov chain  $E_G$  are the  $2^n$  possible arrangements of mutants on the graph. The transition probability between two states  $\mathbf{P}, \mathbf{P}'$  is 0 unless  $|\mathbf{P} \setminus \mathbf{P}'| = 1$  or vice versa. Otherwise, if  $\mathbf{P} \setminus \mathbf{P}' = v^*$ , then the probability of a transition from  $\mathbf{P}$  to  $\mathbf{P}'$  is

$$\frac{\sum_{v \in G \setminus \mathbf{P}} w(v, v^*)}{N + |\mathbf{P}|(r - 1)}$$

where the numerator is the sum of the weights of edges entering  $v^*$  from vertices outside  $\mathbf{P}$ . Similarly, the probability of a transition from  $\mathbf{P}'$  to  $\mathbf{P}$  is

$$\frac{\sum_{v \in \mathbf{P}'} w(v, v^*)}{N + |\mathbf{P}'|(r - 1)}$$

In practice, the resulting matrix is large and not very sparse. Consequently, it can be difficult to work with directly, and we will not revisit it in the course of these notes.)

We now define the notion of  $\rho$ -equivalency.

**Definition 1.** A graph  $G$  is  $\rho$ -equivalent to the Moran process if the cardinality map  $f(\mathbf{P}) = |\mathbf{P}|$  from the states of  $E_G$  to the states of  $M$  preserves the ultimate fixation probabilities of the states. Equivalently, we need

$$\rho(r, G, P, N) = \frac{1 - 1/r^P}{1 - 1/r^N}$$

where  $\rho(r, G, P, N)$  is the probability that a mutant of fitness  $r$  on a graph  $G$ , given any initial population of size  $P$ , eventually reaches the fixation population of  $N$ . (Note that this function is often undefined: on most graphs, different initial conditions with the same number of mutants have different fixation probabilities.)

Note that eq (1) is obtained in the case  $P = 1$ .

This shows that the requirement of preserving fixation probabilities leads inevitably to the preservation of transition probabilities between all the states. In particular, it means that the population size on  $G$ ,  $|\mathbf{P}|$ , performs a random walk with a forward bias of  $r$ , e.g., where the probability of a forward step is  $r/(r + 1)$ .

### **Evolution on circulations is equivalent to the Moran process.**

We now provide a necessary and sufficient condition for  $\rho$ -equivalence to the Moran process for the case of an arbitrary weighted digraph  $G$ . The isothermal theorem for stochastic matrices is obtained as a corollary. First we state the definition of a circulation.

**Definition 2.** The matrix  $W$  defines a circulation  $\leftrightarrow$

$$\forall i, \sum_j w_{ij} = \sum_j w_{ji}$$

This is precisely the statement that the graph  $G_W$  satisfies

$$\forall v \in G, w_o(v) = w_i(v)$$

where  $w_o$  and  $w_i$  represent the sum of the weights entering and leaving  $v$ .

It is now possible to state and prove our first result.

**Theorem 1.** (*Circulation Theorem.*) *The following are equivalent:*

- (1) *G is a circulation.*
- (2)  *$|\mathbf{P}|$  performs a random walk with forward bias  $r$  and absorbing states at  $\{0, N\}$ .*
- (3) *G is  $\rho$ -equivalent to the Moran process*
- (4)

$$\rho(r, G, P, P') = \frac{1 - 1/r^P}{1 - 1/r^{P'}}$$

where  $\rho(r, G, P, P')$  is the probability that a mutant of fitness  $r$  on a graph  $G$  given any initial condition with  $P$  mutants eventually reaches a mutant population of  $P'$ .

*Proof.* We show that (1)  $\rightarrow$  (2)  $\rightarrow$  (3)  $\rightarrow$  (4)  $\rightarrow$  (1).

To see that (1)  $\rightarrow$  (2), let  $\delta_+(\mathbf{P})$  (resp.  $\delta_-(\mathbf{P})$ ) be the probability that the mutant population in a given state increases (resp. decreases), where  $\mathbf{P} \subset \mathbf{V}$  is just the set of vertices occupied by a mutant, corresponding to the present state. The mutant population size will only change if the edge selected in the next round is a member of an edge cut of  $\mathbf{P}$ , e.g., the head is in  $\mathbf{P}$  and the tail is not, or vice-versa.

The probability of a population increase in the next round,  $\delta_+(\mathbf{P})$ , is therefore just the weight of all the edges leaving  $\mathbf{P}$ , adjusted by the fitness of the mutant  $r$ . Thus

$$\delta_+(\mathbf{P}) = \frac{w_o(\mathbf{P})r}{w_o(\mathbf{P})r + w_i(\mathbf{P})}$$

where  $w_o$  and  $w_i$  represent the sum of the weights entering and leaving a vertex set  $\mathbf{P}$ . Similarly,

$$\delta_-(\mathbf{P}) = \frac{w_i(\mathbf{P})}{w_o(\mathbf{P})r + w_i(\mathbf{P})}$$

Dividing, we easily obtain

$$\frac{\delta_+(\mathbf{P})}{\delta_-(\mathbf{P})} = r \frac{w_o(\mathbf{P})}{w_i(\mathbf{P})}$$

We may also observe that

$$\begin{aligned}
w_o(\mathbf{P}) - w_i(\mathbf{P}) &= \left( \sum_{v \in \mathbf{P}} w_o(v) - \sum_{e | e_1, e_2 \in \mathbf{P}} w(e) \right) - \left( \sum_{v \in \mathbf{P}} w_i(v) - \sum_{e | e_1, e_2 \in \mathbf{P}} w(e) \right) \\
&= \left( \sum_{v \in \mathbf{P}} w_o(v) - \sum_{v \in \mathbf{P}} w_i(v) \right)
\end{aligned}$$

where the second and fourth sums in the latter equality are over edges whose two endpoints are in  $\mathbf{P}$ . Since this vanishes when  $G$  is a circulation, we find that on a circulation

$$\forall \mathbf{P} \subset \mathbf{V}, w_o(\mathbf{P}) = w_i(\mathbf{P})$$

and therefore

$$\frac{\delta_+(\mathbf{P})}{\delta_-(\mathbf{P})} = r$$

for all  $\mathbf{P}$ .

Thus the population is simply performing a random walk with forward bias  $r$  as desired, yielding (1)  $\rightarrow$  (2).

(2)  $\rightarrow$  (3) follows immediately from the theory of random walks.

It is easy to see that (3)  $\rightarrow$  (4) by conditional probabilities. We know that

$$\forall P' \geq P, \rho(r, G, P, N) = \rho(r, G, P, P') * \rho(r, G, P', N)$$

Therefore

$$\begin{aligned}
\forall P' \geq P, \rho(r, G, P, P') &= \frac{\rho(r, G, P, N)}{\rho(r, G, P', N)} \\
&= \frac{1 - 1/r^P}{1 - 1/r^N} \left( \frac{1 - 1/r^{P'}}{1 - 1/r^N} \right)^{-1} \\
&= \frac{1 - 1/r^P}{1 - 1/r^{P'}}
\end{aligned}$$

which is the desired result.

To complete the proof, we show that (4)  $\rightarrow$  (1). By (4), we know

$$\rho(1, 2, G, r) = \frac{1 - \frac{1}{r}}{1 - \frac{1}{r^2}} = \frac{r}{r+1}$$

But this is only satisfied for all populations of size 1 if we have

$$\forall v, \frac{\delta_+(v)}{\delta_-(v)} = r$$

As we saw above, this implies that

$$\forall v, w_o(v) = w_i(v)$$

which demonstrates that  $G$  must be a circulation and completes the proof.  $\square$

The isothermal result is just a corollary.

**Theorem 2.** (*Isothermal Theorem.*)  $G$  is  $\rho$ -equivalent to the Moran process  $\Leftrightarrow G$  is isothermal, e.g.,  $W$  is doubly-stochastic.

### Superstars are arbitrarily strong amplifiers of natural selection.

We now sketch the derivation of the amplifier theorem for superstars, denoted  $S_{L,M}^K$ , where  $K$  is the amplification factor,  $L$  the number of leaves, and  $M$  the number of vertices in the reservoir of each leaf. First we must precisely define these objects.

**Definition 3.** The Super-star  $S_{L,M}^K$  consists of a central vertex  $v_{center}$  surrounded by  $L$  leaves. Leaf  $\ell$  contains  $M$  reservoir vertices,  $r_{\ell,m}$  and  $K-2$  ordered chain vertices  $c_{\ell,1}$  through  $c_{\ell,K-2}$ . All directed edges of the form  $(r_{\ell,m}, c_{\ell,1})$ ,  $(c_{\ell,w}, c_{\ell,w+1})$ ,  $(c_{\ell,K-2}, v_{center})$ , and  $(v_{center}, r_{\ell,m})$  exist and no others. In the case  $K = 2$ , the edges are of the form  $(r_{\ell,m}, v_{center})$ , and  $(v_{center}, r_{\ell,m})$ . Illustrations for  $K = 2$  and  $K = 3$  are given in Fig 3. The weight of an edge  $(i,j)$  is given by  $1/d_o(i)$ , where  $d_o$  is the out-degree.

Now we may move on to the theorem.

**Theorem 3.** (*Super-star Theorem.*) *As the number and size of the leaves grows large, the fixation probability of a mutant of fitness  $r$  on a super-star of parameter  $K$  converges toward the behavior of a mutant of fitness  $r^K$  on a circulation:*

$$\lim_{L,M \rightarrow \infty} \rho(S_{L,M}^K) \rightarrow \frac{1 - 1/r^K}{1 - 1/r^{KN}}$$

*Proof. (Sketch)* The proof has several steps.

First we observe that for large  $M$ , the mutant is overwhelmingly likely to appear outside the center or the chain vertices.

Now we show that if the density of mutants in an upstream population is  $d$ , then the probability that an individual in a population immediately downstream will be a mutant at any given time is  $\frac{dr}{1+d(r-1)}$ . In general, if we have  $\eta$  populations, one upstream of the other, the first of which has mutant probability density  $d=d(1)$ , we obtain the following probability density for the  $\nu^{\text{th}}$  population

$$d(\nu) = \frac{dr^\nu}{1 + d(r^\nu - 1)}$$

The result follows inductively from the observation that

$$\begin{aligned} d(j+1) &= \frac{\frac{dr^j}{1+d(r^j-1)}r}{1 + \frac{dr^j}{1+d(r^j-1)}(r-1)} \\ &= \frac{dr^{j+1}}{1 + d(r^{j+1} - 1)} \end{aligned}$$

For the super-star, this result is precise as we move inward from the leaf vertices along the chain leading into the central vertex, where derivation of an analogous result is necessary. Here we require careful bounding of error terms, and allowing  $L$  to go off to the infinite limit. This is in order to ensure that ‘feedback’ is sufficiently attenuated: otherwise, during the time required for information about upstream density to propagate to the central vertex, the upstream population will have already changed too significantly. In this latter regime, ‘memory’ effects can give the resident a very

significant advantage: the initial mutant has died before the central vertex is fully affected by its presence. For sufficiently many leaves feedback is irrelevant to fixation. In the relevant regime we establish that the central vertex is a mutant with probability

$$d(K - 1) = \frac{dr^{K-1}}{1 + d(r^{K-1} - 1)}$$

Our result follows by noting that the probability of an increase in the number of mutant leaf vertices during a given round is very nearly

$$\frac{r}{N + P(r - 1)} \frac{dr^{K-1}}{1 + d(r^{K-1} - 1)} (1 - d)$$

and the probability of a decrease is

$$\frac{1}{N + P(r - 1)} \frac{1 - d}{1 + d(r^{K-1} - 1)} d$$

Dividing, all the terms cancel but an  $r^K$  in the numerator. Thus the mutant population in the leaves performs a random walk with a forward bias of  $r^K$  until fixation is guaranteed or the strain dies out.  $\square$

In the spirit of this result, we may define an amplification factor for any graph  $G$  with  $N$  vertices as the value of  $K$  for which  $\rho(G) = \frac{1-1/r^K}{1-1/r^{KN}}$ . We have seen above that a superstar of parameter  $K$  has an amplification factor of  $K$  as  $N$  grows large.

**The fixation problem for frequency-dependent evolution on graphs is at least as hard as NP.**

NP-hard problems arise naturally in the study of frequency-dependent selection on graphs. Let us consider the general case of some finite number of types; a state of the graph is a partition of its vertices among the types, or a coloring. Given a graph  $G$  and an initial state  $I$ , let VULNERABILITY be the decision problem of whether, given a graph  $G$ , an initial state  $I$ , a small constant  $\epsilon$ , and a desired winning type  $T$ , fewer than  $w$  individuals can mutate to  $T$  so as to ensure fixation of the graph by

T with probability at least  $1-\epsilon$ . By reduction from the Boolean Circuit Satisfiability problem, it can be shown the VULNERABILITY is NP-hard. We omit the details of the proof here.

**Frequency-dependent evolution on graphs leads to a multiplicity of invasion criteria.**

The following computation establishes our observations about directed cycles.

**Proposition 4.** (*Fixation on Directed Cycles.*) *For large  $N$ , the directed cycle favors mutants where  $b > c$  (resp.  $a > d$ ,  $a > c$ ,  $b > d$ ) in the positive symmetric (resp. negative symmetric, positive antisymmetric, negative antisymmetric) orientations.*

*Proof. (Sketch)* For the positive symmetric case, we obtain eq (1) with  $r = b/c$  as a straightforward instance of gambler's ruin with bias  $b/c$ . In the other three orientations, a bit more work is required to account for the case where the patch is of size 1 or size of  $N-1$ . In the negative symmetric and positive antisymmetric orientations, the mutant has an aberrant fitness of  $b$  for patch sizes of exactly 1 (near extinction). In both negative orientations, the resident has an aberrant fitness of  $c$  when the mutant patch is of size  $N-1$  (near fixation). Thus we must do some work to ensure that these aberrations do not ultimately affect which types of mutants are favored on large cycles.

We must evaluate the following expression to obtain the fixation probability of the biased random walk:

$$\rho = \frac{1}{1 + \sum_{i=1}^{N-1} \prod_{j=1}^i \frac{q_j}{p_j}}$$

The values of  $p_i$  and  $q_i$  represent probabilities of increase and decrease when the population is of size  $i$ . We obtain

$$\rho_{-s} = \frac{b(d-a)}{bd - ab - ad + (d/a)^{N-2}(ad + cd - ac)}$$



$$\rho_{+a} = \frac{b(c-a)}{bc - ab - ac + (c/a)^{N-2}(c^2)}$$

$$\rho_{-a} = \frac{b(d-b)}{-b^2 + (d/b)^{N-2}(bd + cd - bc)}$$

for the negative symmetric, positive antisymmetric, and negative antisymmetric cases. For large  $N$ , these expressions are smaller than the neutral fixation probability  $1/N$  if  $d/a$  (resp.  $c/a$ ,  $d/b$ ) is greater than one; if it is less than 1, the fixation probabilities converge to

$$\rho_{-s} = \frac{b(a-d)}{b(a-d) + ad}$$

$$\rho_{+a} = \frac{b(a-c)}{b(a-c) + ac}$$

$$\rho_{-a} = \frac{b(b-d)}{b^2}$$

and the mutant is strongly favored over the neutral case. □

### **Results hold if fertility and mortality are independent Poisson processes.**

Finally, we will make some remarks about our assumptions regarding mutation rate and the meaning of our fitness values.

It is generally the case that suppressing either selection or drift, and in particular the latter, is time intensive. Good amplifiers get arbitrarily large as  $\rho \rightarrow 1$  or 0, and have increasingly significant bottlenecks. Thus, fixation times get extremely long the more effectively drift is suppressed. However, since we are working in the limit where mutations are very rare, this timescale can be ignored. The rate of evolution reduces to the product of population size, mutation rate, and fixation probability.

In our discussions, we have treated fitness as a measure of reproductive fertility. But a range of frequency-independent interpretations of fitness obtain identical results. If

instead of choosing an individual to reproduce in each round with probability proportional to fitness, we choose an individual to die with probability inversely proportional to fitness, and then replace it with a randomly-chosen upstream neighbor, the  $\rho$  values obtained are identical. Put another way, as long as reproduction (leading to death of a neighbor by overcrowding) and mortality (leading to the reproduction of a neighbor that fills the void) are independent Poisson processes, our results will hold.