

# Evolutionary Forces in the Moran Process

Marc Harper<sup>1</sup> and Matteo Smerlak<sup>2,\*</sup>

<sup>1</sup>*Unaffiliated*

<sup>2</sup>*Perimeter Institute for Theoretical Physics,  
31 Caroline St. N., Waterloo ON N2L 2Y5, Canada*

(Dated: September 23, 2016)

## I. INTRODUCTION

Darwinian evolution is by nature a *composite* dynamical process: any evolutionary change involves a combination of selection, mutation and drift. This makes causal inference in evolution much less straightforward, for every evolutionary effect has not one, but at least *three* entangled causes. Delineating the respective importance of these forces in a given evolutionary change would help simplify analysis, but this seems an ill-defined problem: how much of the appearance of a new trait in a population is due to the original mutation? To its subsequent selection through differential reproduction? To drift effects? All three forces are essential components of the evolutionary process.

To make things worse, the selection component is itself of mixed nature, in the sense that it usually acts on several different levels. Examples of this are frequency and density dependence, wherein the fitness effect of a mutation depends not just on the mutant's current genotype, but also on the structure and size of the population to which it belongs. In extreme cases, individual-level and population-level notions of fitness can be opposite: a high-fitness mutant can end up *decreasing* the population mean fitness, as in the prisoner's dilemma. More commonly, the task of assigning relative weights to individual-level vs. population-level selective pressures is hard to define precisely.

The purpose of this paper is to discuss a setting where evolutionary forces actually can be decomposed quantitatively: the Moran birth-death process. The Moran process models evolving multi-type populations in terms of overlapping generations, such that each descendant population is determined stochastically from the ancestral population and pre-determined mutation rates. The population evolves until it stabilizes – fixates in a state of all one type, if mutation is not present,

---

\* matteo.smerlak@gmail.com

or settles into a stationary state of the Markov process, moving between various population states indefinitely.  $\square$

The tool we use to define and quantify the relative strength of evolutionary forces is a general property of Markov processes which (for lack of generally accepted term) we call the *yen*. This concept was originally introduced in the context of non-equilibrium statistical mechanics, where it can be interpreted as a generalization of the notion of heat. More recently, the notion of yen was showed to shed new light on various and seemingly unrelated problems, such as cell replication XXX, wealth inequality XXX. In the context of evolutionary dynamics, yen is closely related to the notion of “fitness flux” discussed in XXX.

An interesting feature of the yen is that it brings to light the *directionality* of Markovian dynamics. In statistical mechanics, this directionality is associated to the second law of thermodynamics, and indeed the concept of yen was invented to formulate generalizations of this law. The implications of such directionality for the Moran evolutionary process will be discussed sec. VII.

## II. THE YEN OF A MARKOVIAN PATH

Consider a Markov chain on a finite set of states  $V$ , characterized by transition probabilities  $T : V \times V \rightarrow [0, 1]$ . The *yen* of a path  $p = (v_1, \dots, v_{|p|}) \in V^n$  with length  $|p|$  is defined as

$$Y(p) = \sum_{i=1}^{|p|-1} \ln \frac{T(v_i, v_{i+1})}{T(v_{i+1}, v_i)} \quad (1)$$

with the convention  $\ln(0/0) = 0$ . Its interpretation is simple: by comparing the forward transition probabilities  $T(v_i, v_{i+1})$  with the backward transition probabilities  $T(v_{i+1}, v_i)$ , the yen measures the system’s “preference” for the path  $p$  over its reverse  $p^{-1} = (v_{|p|}, \dots, v_1)$ . In this sense, the yen generalizes the notion of “landscape” familiar from physics (energy landscape) and biology (fitness landscape) to any Markov chain: if  $v$  and  $w$  are two states of the chain connected by a path  $p$ , we can picture  $v$  as “higher” in the landscape<sup>1</sup> than  $w$  if  $Y(p) < 0$  and vice versa. Note that the yen is trivially odd with respect to reversal,  $Y(p^{-1}) = -Y(p)$ .

The yen has several remarkable properties, some of which are explored in detail in XXX. Here we would like to mention two.

- *Stationary distributions.* Let  $s : V \rightarrow [0, 1]$  be the stationary distribution of the Markov chain, defined equivalently as the limit  $\lim_{k \rightarrow \infty} T^k z$  for any initial distribution  $z : V \rightarrow [0, 1]$

---

<sup>1</sup> Here we use the biologist’s convention for landscape, where peaks are attractive and valleys repulsive; physicists use the opposite convention.

or as the eigenvector of the transition matrix with eigenvalue 1,  $Ts = s$ . (We assume this stationary distribution exists and is unique.) Then for a path  $p$ ,

$$Y(p) > 0 \quad \text{iff} \quad s(v_1) < s(v_{|p|}). \quad (2)$$

This means in particular that we can detect local maxima  $v_*$  of the stationary distribution by the condition that the yen of all length-2 paths starting from  $v_*$  be negative. In this sense the yen can be viewed as a discrete “derivative” of the stationary distribution.

In the special case where the Markov chain is reversible, i.e. satisfies the detailed balance condition  $s(v)T(v, w) = s(w)T(w, v)$  for any pair of states  $(v, w)$ , we have the stronger property that  $Y(p) = \ln s(v_{|p|}) - \ln s(v_1)$  for any path  $p$ . This identity may be interpreted in turn as the Markov-chain analogue of the fundamental theorem of calculus.

- *Entropy.* The entropy of a distribution  $z : V \rightarrow [0, 1]$  is the quantity  $S(z) = -\sum_{v \in V} z(v) \ln z(v)$ . Its growth in time is related to the yen by the inequality<sup>2</sup>

$$S(T^k z) - S(z) \geq \langle -Y(p) \rangle_{k,z} \quad (3)$$

where  $T^k z$  is the evolved distribution after  $k$  time steps and the bracket on the right-hand side denotes an average over all length- $k$  paths  $p = (v_1, \dots, v_k)$  weighted by their probability given the initial distribution  $z$ , viz.  $\text{Prob}(p) = z(v_1)T(v_1, v_2) \cdots T(v_{k-1}, v_k)$ . The inequality (3) is a generalization of the second law of thermodynamics to general Markov chains. It expresses the idea that Markov dynamics is statistically time-asymmetric, with the negative of the yen playing the role of “dissipation” along any given path.

### III. DECOMPOSING EVOLUTIONARY FORCES IN THE MORAN PROCESS

Let us now see how the concept of yen and its properties can be applied to evolutionary dynamics.

#### A. Two population types

##### 1. Two types without mutations

We begin by considering a population with two types  $A$  and  $B$ , with  $a$  individuals of type  $A$  and  $b$  individuals of type  $B$ . We assume that the total population  $a + b = N$  is fixed, so that the

---

<sup>2</sup> This inequality follows from a more general identity known as the *fluctuation theorem* in the physics literature.

population state is entirely determined by the number  $a$  of  $A$ -type individuals. We denote  $\bar{a} = a/N$  and  $\bar{b} = b/N = 1 - \bar{a}$  the corresponding fractions. Finally, we suppose that the fitness of  $A$  and  $B$  individuals are given by two state-dependent functions  $f_A(a)$  and  $f_B(a)$ ; the population mean fitness is  $F(a) = \bar{a}f_A(a) + \bar{b}f_B(a)$ .

The two-type Moran process is then defined as a Markov chain with states  $a \in \{0, \dots, N\}$  and transition probabilities

$$T(a, a+1) = (1 - \bar{a})f_A(a)/F(a), \quad (4)$$

$$T(a, a-1) = \bar{a}f_B(a)/F(a), \quad (5)$$

$$T(a, a) = 1 - T(a, a+1) - T(a, a-1). \quad (6)$$

This Markov chain represents the evolutionary dynamics of overlapping generations: at each time step, one individual is picked at random to die, and is replaced by a new individual with fitness-proportionate probability. It captures the effects of two of the three main forces of evolution, namely selection and drift.

Consider now the yen of a transition  $a \mapsto a+1$ , i.e.  $Y(a, a+1) = \ln[T(a, a+1)/T(a+1, a)]$ . According to the definition of the transition probabilities above,  $Y(a, a+1)$  splits into three terms:

$$Y(a, a+1) = \ln \frac{f_A(a)}{f_B(a+1)} + \ln \frac{F(a+1)}{F(a)} + \ln \frac{1 - \bar{a}}{(a+1)}. \quad (7)$$

Each term captures a different component of the evolutionary change  $a \mapsto a+1$ . The first term  $Y_s$  depends on the relative fitness of  $A$  and  $B$ , taking positive values when  $A$  is fitter than  $B$ : it therefore measures the contribution of *natural selection* of  $A$  over  $B$  in the transition  $a \mapsto a+1$ . The second term measures  $Y_a$  the growth of the population mean fitness in the descendant population compared to the ancestral population; in this sense we can say that  $Y_a$  expresses the role of *collective adaptation* in favoring the transition  $a \mapsto a+1$ . Finally the third term  $Y_d$  is independent of fitness and expresses the relative likelihood of choosing an individual of type  $A$  or  $B$  for death: it is the contribution of *drift* to the yen of  $a \mapsto a+1$ .

## 2. Two types with mutations

The two-type Moran process can be generalized to include mutations. To this effect we introduce mutation rates  $\mu_{AB}(a)$  and  $\mu_{BA}(a)$ , giving the probabilities that a type- $A$  individual mutates into a type- $B$  individual just before being introduced into the population (and reciprocally). Typical mutation rates are  $\mathcal{O}(1/N)$ .

In the presence of mutations the transition probabilities become

$$T(a, a+1) = (1 - \bar{a}) \frac{f_A(a)[1 - \mu_{AB}(a)] + f_B(a)\mu_{BA}(a)}{F(a)}, \quad (8)$$

$$T(a, a-1) = \bar{a} \frac{f_B(a)[1 - \mu_{BA}(a)] + f_A(a)\mu_{AB}(a)}{F(a)}, \quad (9)$$

$$T(a, a) = 1 - T(a, a+1) - T(a, a-1). \quad (10)$$

Assuming that  $\mu_{AB} = \mu = \mu_{BA}$ , we can then decompose the yen into components for each of the evolutionary processes, using a Taylor expansion in  $\mu$ :

$$\begin{aligned} \sigma(\mathbf{a} \rightarrow \mathbf{a} + \mathbf{1}) = & \underbrace{\log \frac{\bar{f}_A(\mathbf{a} + \mathbf{1})}{\bar{f}_B(\mathbf{a})} \frac{f_A(\mathbf{a})}{f_B(\mathbf{a} + \mathbf{1})}}_{\text{Fitness}} + \underbrace{\log \frac{(N - a - 1)(a + 1)}{(N - a)(a)}}_{\text{Drift}} \\ & + \underbrace{\mu \left( \frac{N - a}{a} \frac{f_B(\mathbf{a})}{f_A(\mathbf{a})} - \frac{a + 1}{N - a - 1} \frac{f_B(\mathbf{a} + \mathbf{1})}{f_A(\mathbf{a} + \mathbf{1})} \right)}_{\text{Mutation}} \end{aligned}$$

This decomposition gives us the individual contributions of each of the evolutionary forces on the dynamics of the Markov process.

### 3. Examples

As a first example, consider the simple case where fitness is state-independent, i.e. without any density or frequency dependent effects. Also assume for simplicity that  $f_A = 1$  ( $A$  is the “wild type”), and denote  $f_B = 1 + s$  the fitness of  $B$  (the “mutants”), with  $s \ll 1$  a small selection coefficient. In this case the yen reduces to

$$Y(a, a+1) \simeq s + \ln \frac{N - a}{a + 1}. \quad (11)$$

This expression is compatible with the intuitive expectation that evolution has two distinct regimes: when  $a \simeq N/2$  (wild type and mutant), the drift component becomes negligible and the yen reduces to the selection coefficient ( $Y \simeq s$ ), i.e. *selection* is the dominant evolutionary force; for  $a \simeq 0$  or  $a \simeq N$ , the selection component is negligible in the yen ( $Y \simeq -\ln \bar{a}$  and  $Y \simeq \ln(1 - \bar{a})$  respectively), and *drift* dominates evolution. In this simple setting collective adaptation of course plays no role.

Another example of a two-type Moran process is the Hawk-Dove game, in which the fitness landscape  $f(a) = (f_A(a), f_B(a))$  can be written as

$$f(a) = \begin{pmatrix} 1 & q \\ q & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}, \quad (12)$$

where  $q > 1$  is a fixed parameter.

For a concrete example, consider a fitness landscape given by a Hawk-Dove game matrix,

$$f(\mathbf{a}) = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \mathbf{a},$$

we can plot the relative contributes of each evolutionary process (see Figure ?? 3.png). When the  $\mu$  is zero the stationary distribution has a local maximum at the center state  $(N/2, N/2)$  if  $N$  is even, which is the evolutionarily stable state for the landscape. Drift is largest (in absolute value) near the boundary, as are the other components, with adaptation being the smallest typically. When the game matrix is altered to another member of the Hawk-Dove class (4.png):

$$f(\mathbf{a}) = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} \mathbf{a},$$

the  $\mu$  is again zero at the stationary maximum (approximately  $a = 12$ ), and the relative fitness becomes a more dominant influence on the left boundary. With a tournament selection matrix (0.png), mutation becomes by far the most influential factor on the right boundary. Mutation also drives the neutral landscape in from the boundary (1.png).

The classical Moran process has fitness landscape of the form:

$$f(\mathbf{a}) = \begin{pmatrix} r & r \\ 1 & 1 \end{pmatrix} \mathbf{a},$$

where  $r$  is the relative fitness of the first type. Figure ?? (2.png) shows the decomposition in the case of  $r = 2$  and  $N = 20$ . For mutation rate  $\mu = 0$  the process would fixate into the two states  $(N, 0)$  and  $(0, N)$  with a probability depending on  $r$ . For nonzero  $\mu = 1/N$  we obtain the stationary distribution in Figure ??. As the population evolves, states with more individuals of type  $A$  are favored, and we see from the decomposition that mutation and relative fitness dominate the dynamic, with drift playing a smaller role as the population approaches the boundary. For larger populations, drift typically plays a diminished role, as expected.

## B. General Moran process

We focus on the Moran process with mutation, as formulated in [? ]; see also [? ], [? ], and [? ]. Let a population of size  $N$  be composed of  $n$  types  $A_1, \dots, A_n$  with  $a_i$  individuals of type  $A_i$  so that  $N = a_1 + \dots + a_n$ . We denote a population state by the tuple  $\mathbf{a} = (a_1, \dots, a_n)$  and the

population distribution by  $\bar{\mathbf{a}} = \mathbf{a}/N$ . Define a matrix of mutations  $M$ , where  $0 \leq M_{ij}(\mathbf{a}) \leq 1$  may be a function of the population state in general, but we will typically assume in examples that for some constant value  $\mu$ , the mutation matrix takes the form  $M_{ij} = \mu/(n-1)$  for  $i \neq j$  and  $M_{ii} = 1 - \mu$ . A typical mutation rate is  $\mu \approx 1/N$ .

The Moran process with mutation is a Markov chain on the population states defined by the following transition probabilities, corresponding to a birth-death process where birth is fitness-proportionate (with mutation) and death is uniformly random. To define the adjacent population states, let  $\mathbf{i}_{\alpha,\beta}$  be the vector that is 1 at index  $\alpha$ ,  $-1$  at index  $\beta$ , and zero otherwise, with the convention that  $\mathbf{i}_{\alpha\alpha}$  is the zero vector of length  $n$ . Every adjacent state of state  $\mathbf{a}$  for the Moran process is of the form  $\mathbf{a} + \mathbf{i}_{\alpha,\beta}$  for some  $1 \leq \alpha, \beta \leq n$ —one more type  $\alpha$  individual, one less type  $\beta$  individual.

At a population state  $\mathbf{a}$  we randomly choose an individual of type  $A_i$  to reproduce proportionally to its fitness  $f_i(\bar{\mathbf{a}})$ , allowing for mutation of the new individual as given by the mutation probabilities. The distribution of fitness proportionate selection probabilities is given by  $p(\bar{\mathbf{a}}) = M(\bar{\mathbf{a}})\bar{\varphi}(\bar{\mathbf{a}})$ ; explicitly, the  $i$ -th component is

$$p_i(\bar{\mathbf{a}}) = \frac{\sum_{k=1}^n \varphi_k(\bar{\mathbf{a}}) M_{ki}}{\sum_{k=1}^n \varphi_k(\bar{\mathbf{a}})} \quad (13)$$

where the function  $\varphi(\bar{\mathbf{a}}) = \bar{\mathbf{a}}_i f_i(\bar{\mathbf{a}})$ . This yields the transition probabilities

$$\begin{aligned} T(\mathbf{a}, \mathbf{a} + \mathbf{i}_{\alpha,\beta}) &= p_\alpha(\bar{\mathbf{a}}) \bar{\mathbf{a}}_\beta \quad \text{for } \alpha \neq \beta \\ T(\mathbf{a}, \mathbf{a}) &= 1 - \sum_{\mathbf{b} \text{ adj } \mathbf{a}, \mathbf{b} \neq \mathbf{a}} T(\mathbf{a}, \mathbf{b}) \end{aligned} \quad (14)$$

#### IV. THREE POPULATION TYPES

For populations of more than two types, there are more directions in which the process can move, and so more yens to consider. Fixing the two components which are changing, e.g.  $(a, b, c) \rightarrow (a+1, b, c-1)$  for a population of three types  $A$ ,  $B$ , and  $C$  allows the yen to be decomposed as for two type populations, with the appropriate substitutions of e.g.  $f_C$  for  $f_B$  and the population states. Instead of one direction for yen (+ or -) in a two type population, in a three type population there are three possible directions.

Consider the fitness landscape given by

$$f(\mathbf{a}) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \mathbf{a},$$

and the Moran process with  $\mu = 3/(2N)$  and strength of selection  $\beta = 0.1$ . For simplicity we assume that  $N$  is divisible by 3. Then we have several local maxima of the stationary distribution: one interior:  $(N/3, N/3, N/3)$ , three on the boundary:  $(0, N/2, N/2)$ ,  $(N/2, 0, N/2)$ ,  $(N/2, N/2, 0)$ , and the three corner points:  $(0, 0, N)$ ,  $(0, N, 0)$ ,  $(N, 0, 0)$ . Although the process is not reversible and a closed-form of the stationary distribution can not be easily found (if at all), we can simply check which states have the property that the maximum of the incoming yens is negative. These all correspond to the local maxima of the stationary distribution (see ??). Moreover, we see from the max incoming yen plot that there is an “inner boundary” that is rarely traversed, and that the corner points are relatively isolated, with only mutation to kick the trajectory out. Finally, for  $N = 60$ , three local minima at  $(3, 3, 54)$ ,  $(3, 54, 3)$ ,  $(54, 3, 3)$  are (properly) detected by the maximum incoming yen.

## V. POPULATION DYNAMICS ON GRAPHS

Variants of the Moran process where the individuals occupy vertices on a (directed) graph are popular. The introduction of a graph is known to alter the fixation probabilities, strength of selection, and other properties of the process [?] [?], and we will now see that the graph structure alters the properties of the stationary distribution and the relative strengths of evolutionary forces. Typically two dynamics are studied, *birth-death* where an individual is selected globally in a fitness proportionate manner to replace (possibly with mutation) one of its outgoing neighbors selected at random; or *death-birth*, where an individual is chosen uniformly at random to be replaced and one of the incoming neighbors is selected to replace proportionally to the fitness of the neighbors. A complete graph with birth-death dynamics would correspond to the Moran process with mutation described above.

Adding the graph structure typically greatly increases the number of states of the Markov process. For example, for an undirected cycle and a population of two types of size  $N$  there are  $2^N$  states corresponding to all the possible distributions of individuals about the cycle (two in each state). If the graph is not fixed in space, the state space can be reduced by the symmetry group of the cycle (the dihedral group) of order  $2N$  to  $2^{N-1}/N$ . Regardless, the state space is exponentially larger than that of the classical Moran process which has  $N + 1$  states. Because the state space becomes large quickly, the stationary distribution is difficult to study computationally, and before now, analytically. The yen gives us a tool to discover the local extrema without having to compute the stationary distribution directly.



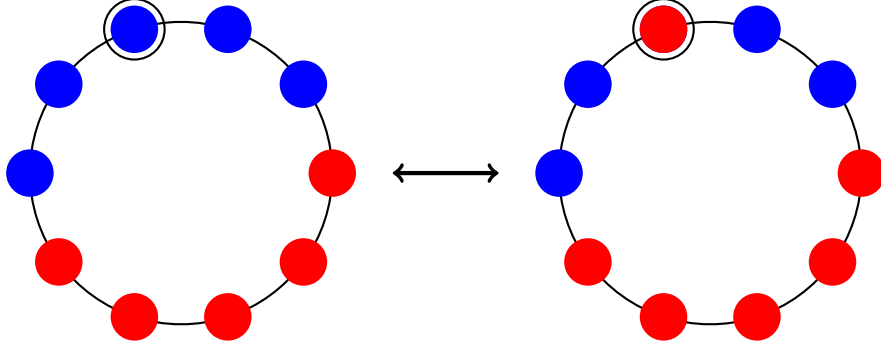


FIG. 1. Transition showing two connected subpopulations where one population is separated by a mutation.

As there are now many more ways in which the population can achieve a given mix of types, the most being  $\binom{2m}{m}$  for a population distribution of  $(m, m)$ , there are potentially many more local maxima and minima. In fact, an  $(m, m)$  population distribution can be both a local maximum, when the populations occupy two adjacent semi-circles of the cycle, and a local minimum, when the individuals are of alternating type about the cycle. The yen tells us that in first case a small mutation rate is a stabilizing force whereas in the second case a small mutation rate leads to instability.

Let us define the transition probabilities between states on the cycle in the two player case. At a given graph configuration with population distribution  $(a, b)$ , we choose an individual to reproduce (with mutation) globally. This does not change the selection ratio from the classical case. Then we choose one of the neighbors to replace, and now it matters how the population is distributed on the graph. The configuration may not change at all, if the replaced individual is of the same type as the outcome of the selection-mutation step. For example, if the population state is  $(m, m)$  where the individuals are distributed as two connected semi-circles, and individual selected for reproduction inside one of the two semi-circles will only change the graph configuration if a mutation occurs. However for the four vertices where the subpopulations are adjacent, replication without mutation and into the other subpopulation will change the configuration.

With the yen we can show that for the Hawk-Dove fitness landscape that this population state is a local maximum when mutation is small, and that for the neutral fitness landscape that there is a plateau of states connected to this one that form a collective extrema. We need to show that the yen of each possible transition is negative. Suppose  $(a, b) = (m, m)$ ,  $m > 4$ , and  $\mu < 1/2$  and consider the transition depicted in Figure 1. The transitions probabilities between the two configurations are given by (left-to-right first):

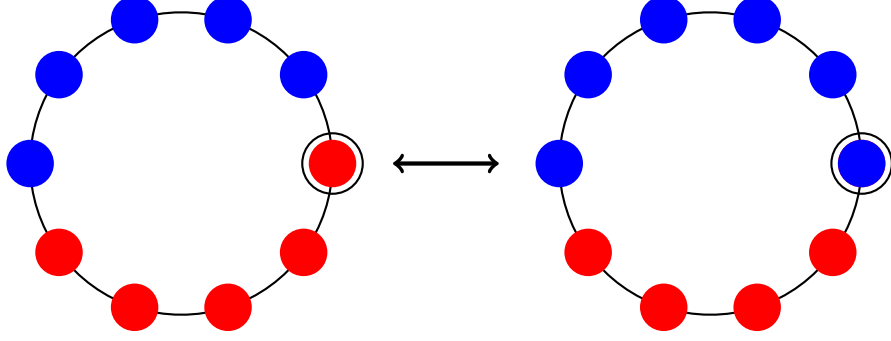


FIG. 2. Transition showing two connected subpopulations where one population is extends into the other without a mutation.

$$T = \underbrace{\frac{a f_A(a)}{\bar{f}(a)} \frac{2}{a}}_{\text{Select a neighboring } A} \underbrace{\frac{1}{2}}_{\text{Replace inward}} \underbrace{\mu}_{\text{mutate}}$$

$$T' = \underbrace{\frac{(a-1) f_A(a-1)}{\bar{f}(a-1)} \frac{2}{a-1}}_{\text{Select a neighboring } A} \underbrace{\frac{1}{2}}_{\text{Replace inward}} \underbrace{(1-\mu)}_{\text{Don't mutate}}$$

The yen is then given by

$$\log \frac{T}{T'} = \underbrace{\log \frac{\mu}{1-\mu}}_{\text{Mutation}} + \underbrace{\log \frac{f_A(a)}{f_A(a-1)}}_{\text{Selection}} + \underbrace{\log \frac{\bar{f}(a-1)}{\bar{f}(a)}}_{\text{Adaptation}}$$

In most cases, this yen is less than zero: for the neutral fitness landscape, for the Hawk-Dove landscape, for continuous landscapes and large  $N = 2m$  or small  $\mu$ . In contrast to the yen decomposition for the well-mixed case above where selection is typically the dominating force at interior states, here we see that mutation is the dominating force due to the introduction the graph structure. A small mutation rate yields a large negative term that washes out the variation in a smooth landscape except in extreme cases (very small  $N$  or very sharp changes in the landscape).

The analogous transition within the  $B$  population is similar, so we move on to the four transitions where one population elongates into other. It will suffice to consider one such transition with  $A$  extending into  $B$ , which can occur in two ways: either the  $A$  subpopulation extends without mutation or the  $B$  population shortens due to mutation as depicted in Figure 2. The extending transition probability is given by:

$$\frac{1}{a} \frac{a f_A(a)}{\bar{f}(a)} \frac{1}{2} (1-\mu) + \frac{1}{b} \frac{b f_B(a)}{\bar{f}(a)} \frac{1}{2} \mu = \frac{1}{2} \frac{(1-\mu) f_A(a) + \mu f_B(a)}{\bar{f}(a)}$$

The reverse transition probability is given by:

$$\frac{1}{b-1} \frac{(b-1)f_B(a+1)}{\bar{f}(a+1)} \frac{1}{2}(1-\mu) + \frac{1}{a+1} \frac{(a+1)f_A(a+1)}{\bar{f}(a+1)} \frac{1}{2}\mu = \frac{1}{2} \frac{(1-\mu)f_B(a+1) + \mu f_A(a+1)}{\bar{f}(a+1)}$$

Finally, the yen can be written as

$$Y = \log \frac{(1-\mu)f_A(a) + \mu f_B(a)}{(1-\mu)f_B(a+1) + \mu f_A(a+1)} + \log \frac{\bar{f}(a+1)}{\bar{f}(a)}$$

Here we have given an exact expression for all  $\mu$ ; a Taylor expansion would produce a decomposition similar to the well-mixed case. The fitness landscape now largely controls the value of the yen when  $\mu$  is small since the first term will reduce to a relative fitness term  $f_A(a)/f_B(a+1)$ , similar to the behavior seen in the well-mixed model. For the Hawk-Dove landscape, we have that the yen at the state with  $a = m = b = N/2$  becomes

$$\log \frac{\frac{3N}{2}}{\frac{3N}{2} - 1 + 2\mu} + \log \frac{\frac{3N^2}{2} - 1}{\frac{3N^2}{2}},$$

which is less than zero if  $\mu < (1/2)(1+1/N)$ , as is typical. Combined with the cases above, we have shown that the state with two equally-sized and connected subpopulations is a local maximum of the stationary distribution, as are the  $N$  states that are rotations of this state.

For the neutral fitness landscape (eliminating the contribution of the forces of selection and adaptation), this yen is zero, which implies that these two states have equal stationary probability. This means that there is a subset of approximately  $N^2$  states (consisting each of two connected subpopulations of any size up to rotation) of equal stationary probability, in contrast to the classical case in which the  $(m, m)$  central state is a local maximum when e.g.  $\mu = 1/N$ .

Similar reasoning applies to other landscapes, and to the states in which there are four equally-sized subpopulations each occupying a contiguous quarter of cycle (of which there are  $N/2$  rotations), and so on. This means that for Hawk-Dove there are many interior local maxima (rather than one in the classical case). For a population of size  $N$  divisible by  $2^k$ , there are  $\approx 2N$  such maxima (not counting states that do not fit this pattern):

$$\sum_{i=1}^{k-1} \frac{N}{2^{i-1}} \approx 2N.$$

By introduction a graph structure to the population, we have altered the relative strengths of the evolutionary forces because the population dynamic now moves between many highly structured states rather than a much smaller collection of amorphous states. Drift is absent from the transitions discussed above when we focus on the movement between two specific states that have

a lot of uniformity; this is because in many cases the selection of the individual to reproduce affects the probabilities of which type of individual will be replaced. Specifically, if the individuals are clustered by type, then we are more likely to choose an individual of the same type to be replaced than if the population were to be well-mixed. On the other hand, in the case where the individuals are alternating about the cycle, drift becomes a dominant force, particularly when the (global) fitness of each type is equal: to which state the population transitions then depends only on which individual is randomly chosen to reproduce and whether a mutation occurs or not. Now we are more like to choose an individual of the opposing type to be replaced, and if a mutation does not occur, the configuration will change. Hence it should now come as no surprise that the changes in the relative strengths of the evolutionary forces due to the cycle structure induce changes in the stationary distribution, and we have shown that the stationary distribution indeed has many more extrema.

The yen analysis of extrema easily shows that certain states are local minima. For example, the configurations where the two types alternate are typically local minima as any replication event without a mutation will leave this state, but returning to the state is not as easy. Computations indicate that there can be many extrema detected by the yen depending on the landscape and mutation rate: in the case a population of size  $N = 12$  with the Hawk-Dove landscape and a mutation rate of  $\mu = 3/\sqrt{N}$ , there are approximately 200 extrema, mostly minima, out of  $2^N = 2046$  configurations. We also note that the stationary distribution is very skewed. For the same process with  $\mu = 1/N$ , there are 72 local maxima, 56 local minima, and 134 of 4096 states have 56% of the stationary probability (each with at least 0.1 %). Approximately 2950 states have stationary probability less than 0.01 %. The stationary distribution for the neutral landscape is skewed in the same manner.

This entire analysis extends to  $k$ -regular graphs and fitness landscapes with equilibria at  $(a, b)$  where  $a \neq b$ . We can also study the death-birth case similarly. Finally, we expect similar results for populations with more than two types, with cycles segregated into subpopulations of approximately three equal sizes being local maxima for the 3x3 landscape considered above.

## VI. YEN AS A SEARCH ALGORITHM

We can use the fact that at a local maximum the yen is positive for each neighboring state (and negative for minima) to formulate a search algorithm for local extrema of stationary distributions. This is particularly useful in the cases like the graph process above, where the number of states

is enormous. As long as the list of neighbors and the transition probabilities can be computed as needed, we can proceed as follows (psuedocode) to search for extrema:

Assume a given initial state  $x$

While True:

    Compute the set of neighboring states  $z_i$

    Compute the transitions  $T(x, z_i), T(z_i, x)$

    Compute the yens  $y_i = \log (T(x, z_i) / T(z_i, x))$

    If all the yens are positive:

        break with  $x$  as a local maximum

    Else:

        chose  $z_i$  with smallest zen, if not yet visited (else return  $x$ )

        set  $x = z_i$

Note that if there is a subset of states with equal stationary probabilities that is collectively maximal or minimal the algorithm will return one state of the subset.

We implemented this algorithm in Python along with the graph process described in the preceeding section. As expected, the algorithm finds the expected maximum even when the population size is large ( $N = 512$ ) and the state space is of size  $2^5 11/512$  (given typical parameters and starting space). Convergence is fast even with this basic algorithm. [https://github.com/marcharper/Yen/blob/master/yen\\_search.py](https://github.com/marcharper/Yen/blob/master/yen_search.py)

## VII. DIRECTIONALITY OF EVOLUTION

Second law and relationship with the germans