Discrete evolutionary genetics. Multiplicative fitnesses and the mutation-fitness balance

Thierry Huillet, Servet Martinez

DISCRETE EVOLUTIONARY GENETICS. MULTIPLICATIVE FITNESSES AND THE MUTATION-FITNESS BALANCE

THIERRY HUILLET[1], SERVET MARTINEZ[2]

Abstract. We revisit the multi-allelic mutation-fitness balance problem especially when fitnesses are multiplicative. Using ideas arising from quasi-stationary distributions, we analyze the qualitative differences between the fitness-first and mutation-first models, under various schemes of the mutation pattern. We give some stochastic domination relations between the equilibrium states resulting from these models.

Keywords: Evolutionary genetics, fitness landscape, selection, mutation, stochastic models, quasi-stationarity.

AMS 60J70, 92A10

PACS numbers: 02.50.Ey, 87.23.Kg, 87.10.Mn

1. Introduction and outline

Population genetics aims at elucidating the fate of the allelic population composition when various driving 'forces' such as selection or mutation are at stake in the gene pool. This requires to identify first the updating mechanisms responsible of the gene frequency-distributions evolution over time. In this note, we shall briefly revisit the basics of the deterministic dynamics arising in discrete-time asexual evolutionary genetics when the origin of motion is either the fitness or the mutations or both. We start with the multi-allelic haploid case before dealing with the diploid case. First, we consider general fitness mechanisms, then general mutation mechanisms and then we shall combine the two.

The general purpose of the Sections $2-3$ is to introduce separately the marginal allelic dynamics driven by fitness and then the one driven by mutations. These issues are of course part of the standard models discussed for example in [1], [2], [3] and [4].

In Section 4, we stress that there are two different ways to combine the fitness and the mutation effects. One (fitness-first), which is classical, consists in applying first the fitness mapping and then let mutation act on the result. The other (mutation-first) consists in reversing the order. Stochastic models pertaining to the mutation/selection combination are numerous. See [5], [6] (and the References therein) for the relation of a mutation/selection model with ancestral branching processes. A recent discussion on a Markov chain evolution to study the probability that a new mutant becomes fixed in a Moran type model can be found in [7]. A work describing phenotypic variation and natural selection by modeling population as a Markov point process can be found in [8].

In Section 5, we focus on a model with multiplicative fitnesses and general mutation pattern and we analyze both the fitness-first and mutation-first dynamics. Starting with the fitness-first dynamics, we observe that it has the structure of a discrete-time nonlinear master equation of some Markov process whose construction we give. In this stochastic interpretation, the polymorphic equilibrium state interprets as a quasi-stationary distribution of the Markov process conditioned to be currently alive. It is the left eigenvector of some sub-stochastic matrix A associated to its spectral radius. The corresponding right survival eigenvector makes sense in this interpretation. A similar interpretation can be given when dealing with the mutation-first dynamics driven now by some sub-stochastic matrix B with its own left and right Perron-Frobenius eigenvectors. The matrices A and B are diagonally similar. Using these stochastic tools, we observe that the mean fitness at equilibrium of the model B is larger than the one of model A, together with some stochastic domination properties between both the left and right Perron-Frobenius eigenvectors of the models A and B. If we specify the structure of the mutation matrix to be reversible, then the right and left Perron-Frobenius eigenvectors of each model can be related to one another by using an appropriate Schur product. Some simplifications also occur if we deal with symmetric mutations because the right (left) eigenvector of A coincides with the left (right) eigenvector of B.

Section 6 particularizes the study of Section 5 when a house of cards condition holds for the mutation matrix. Because this mutation model is quite restrictive, some simplifications occur and the shapes of the polymorphic equilibrium states can be made more explicit.

The interpretation of the fitness-first and mutation-first dynamics in terms of a stochastic process conditioned on not being currently absorbed in some coffin state suggests that related conditional models for the evolutionary dynamics involving multiplicative fitness and mutations could also be relevant. In Section 7, we suggest to condition the process on its non-extinction either locally (stepwise) or to condition it globally on not getting extinct in the remote future. Models A and B lead to different conditional dynamics.

2. Evolution under fitness: the deterministic point of view

We briefly describe the frequency distribution dynamics when fitness only drives the process. We start with the haploid case before moving to the diploid case.

2.1. Single locus: haploid population with K alleles.

Consider K alleles A_k, $k = 1, ..., K$ attached to a single locus. Suppose the current time-t allelic frequency distribution is given by the column vector $\mathbf{x} := x_k$, $k = 1, ..., K$ [1]. We therefore have $\mathbf{x} \in S_K = \left\{ \mathbf{x} \geq \mathbf{0} : \ |\mathbf{x}| := \sum_{k=1}^{K} x_k = 1 \right\}$ the $K-$simplex. Let $\mathbf{w} := w_k > 0$, $k = 1, ..., K$, denote the absolute fitnesses of the alleles. Let

$$w(\mathbf{x}) := \sum_l w_l x_l = \mathbf{w}^* \mathbf{x} \tag{1}$$

[1]In the sequel, a boldface variable, say $\mathbf{x}$, will represent a column-vector so that its transpose, say $\mathbf{x}^*$, will be a line-vector. Similarly, A^* will stand for the transpose of some matrix A.

be the mean fitness of the population at time t. The variance in absolute fitness $\sigma^2(\mathbf{x})$ and the variance in relative fitness $\overline{\sigma}^2(\mathbf{x})$ are given respectively by

$$\sigma^2(\mathbf{x}) = \sum_{k=1}^{K} x_k \left(w_k - w(\mathbf{x})\right)^2 ; \, \overline{\sigma}^2(\mathbf{x}) = \sum_{k=1}^{K} x_k \left(\frac{w_k}{w(\mathbf{x})} - 1\right)^2 = \frac{\sigma^2(\mathbf{x})}{w(\mathbf{x})^2}. \tag{2}$$

2.1.1. *Dynamics.* The discrete-time update of the allele frequency distribution on the simplex S_K is given by [2]:

$$x'_k = p_k(\mathbf{x}) := \frac{x_k w_k}{w(\mathbf{x})}, \; k = 1, ..., K. \tag{3}$$

As required, the vector $\mathbf{p}(\mathbf{x}) := p_k(\mathbf{x})$, $k = 1, ..., K$, maps S_K into S_K. In vector form, with $D_{\mathbf{x}} :=$diag$(x_k, k = 1, ..., K)$, the nonlinear deterministic dynamics reads [3]:

$$\mathbf{x}' = \mathbf{p}(\mathbf{x}) = \frac{1}{w(\mathbf{x})} D_{\mathbf{w}} \mathbf{x} = \frac{1}{w(\mathbf{x})} D_{\mathbf{x}} \mathbf{w},$$

or, with $\Delta\mathbf{x} := \mathbf{x}' - \mathbf{x}$, the increment of $\mathbf{x}$

$$\Delta\mathbf{x} = \left(\frac{1}{w(\mathbf{x})} D_{\mathbf{w}} - I\right) \mathbf{x}.$$

Without loss of generality, we can assume that $0 < w_1 \leq ... \leq w_K = 1$. Thus that allele A_K has largest fitness.

Let $\varphi_{k,l} : S_K \to S_K$ be the involution exchanging the coordinates k and l. When $w_k = w_l$, we have that $\varphi_{k,l}(\mathbf{x}') = \left(\varphi_{k,l}(\mathbf{x})\right)'$, and so the evolution is symmetric under $\varphi_{k,l}$. In that case, the alleles k and l can be merged into a single one.

For $\mathbf{x} \in S_K$, let support$(\mathbf{x}) = \{k : x_k > 0\}$. Let $\mathcal{E} = \{k : w_k = 1\}$ be the set of alleles with maximal fitness. Any $\mathbf{x} \in S_K$ such that support$(\mathbf{x}) \subseteq \mathcal{E}$ is called an equilibrium state. A vector $\mathbf{x} = \mathbf{e}_k = (0, ..., 0, 1, 0, ..., 0)$ with $k \in \mathcal{E}$ is called a pure (or monomorphic) equilibrium state.

2.1.2. *Mean fitness increase.* According to the dynamical system (3), unless its equilibrium state is attained, the absolute mean fitness $w(\mathbf{x})$ increases. Indeed, with $\Delta w(\mathbf{x}) = w(\mathbf{x}') - w(\mathbf{x})$:

$$\Delta w(\mathbf{x}) = \sum_k w_k \Delta x_k = \sum_k w_k x_k \left(\frac{w_k}{w(\mathbf{x})} - 1\right) = \frac{\sum_k w_k^2 x_k}{w(\mathbf{x})} - w(\mathbf{x}) \geq 0, \tag{4}$$

and it is > 0 except when support$(\mathbf{x}) \subseteq \mathcal{E}$.

The mean fitness is maximal at equilibrium. The rate of increase of $w(\mathbf{x})$ is:

$$\frac{\Delta w(\mathbf{x})}{w(\mathbf{x})} = \sum_k x_k \left(\frac{w_k}{w(\mathbf{x})} - 1\right)^2 = \sum_k \frac{(\Delta x_k)^2}{x_k} = \overline{\sigma}^2(\mathbf{x}). \tag{5}$$

These last two facts are sometimes termed the 1930s Fisher fundamental theorem of natural selection (FTNS). Then, if there is an allele whose fitness is strictly larger than the ones of the others starting from any initial state of S_K which is not an

[2]The symbol $'$ is a common and useful notation to denote the updated frequency.

[3]$D_{\mathbf{x}}\mathbf{w}$ clearly is the Schur product of $\mathbf{x}$ and $\mathbf{w}$. See [3] page 238 for a similar notational convenience.

extremal point, the haploid trajectories will converge to this fittest state.

2.2. **Single locus: diploid population with K alleles.** We now run into similar considerations but with diploid populations.

2.2.1. *Joint evolutionary dynamics.* Let $w_{k,l} \geq 0,\ k,l = 1,...,K$ stand for the absolute fitness of the genotypes A_kA_l attached to a single locus. Assume $w_{k,l} = w_{l,k}$ ($w_{k,l}$ being proportional to the probability of an A_kA_l surviving to maturity, it is natural to take $w_{k,l} = w_{l,k}$). Let then W be the symmetric fitness matrix with $k,l-$entry $w_{k,l}$.

Assume the current frequency distribution at time t of the genotypes A_kA_l is given by $x_{k,l}$. Let X be the frequencies array with $k,l-$entry $x_{k,l}$. The joint evolutionary dynamics in the diploid case is given by the updating:

$$x'_{k,l} = x_{k,l}\frac{w_{k,l}}{\omega(X)} \text{ where } \omega(X) = \sum_{k,l} x_{k,l}w_{k,l}. \tag{6}$$

The relative fitness of the genotype A_kA_l is $w_{k,l}/\omega\left(X\right)$. The joint dynamics takes the matrix form:

$$X' = \frac{1}{\omega(X)}X \circ W = \frac{1}{\omega(X)}W \circ X$$

where $\circ$ stands for the (commutative) Hadamard product of matrices.

Let J be the $K \times K$ matrix whose entries are all 1 (the identity for $\circ$). Then

$$\Delta X := X' - X = \frac{1}{\omega\left(X\right)}\left(X - J\right) \circ W = \frac{1}{\omega\left(X\right)}W \circ \left(X - J\right).$$

Let

$$\sigma^2(X) = \sum_{k,l=1}^{K} x_{k,l}(w_{k,l} - \omega(X))^2;\ \overline{\sigma}^2(X) = \sum_{k,l=1}^{K} x_{k,l}\left(\frac{\omega_{k,l}}{\omega(X)} - 1\right)^2 = \frac{\sigma^2(X)}{\omega(X)^2} \tag{7}$$

stand respectively the genotypic variance in absolute fitness and the diploid variance in relative fitness. The increase of the mean fitness is given by

$$\Delta\omega\left(X\right) = \sum_{k,l}\Delta x_{k,l}w_{k,l} = \sum_{k,l} x_{k,l}\left(\frac{w_{k,l}^2}{\omega(X)} - w_{k,l}\right) = \omega(X)\overline{\sigma}^2(X) \geq 0, \tag{8}$$

which vanishes only at the equilibrium states maximizing $w_{k,l}$, with a relative rate of increase: $\Delta w(X)/w(X) = \overline{\sigma}^2(X)$. This is the diploid version of the FTNS.

2.2.2. *Marginal allelic dynamics.* Assuming a Hardy-Weinberg equilibrium, the frequency distribution at time t, say $x_{k,l}$, of the genotypes A_kA_l is given by: $x_{k,l} = x_kx_l$ where $x_k = \sum_l x_{k,l}$ is the marginal frequency of allele A_k in the whole genotypic population. The frequency information is $\mathbf{x} = X\mathbf{1}$ ($\mathbf{1}$ is the unit K-vector) and the mean fitness is given by the quadratic form: $\omega(\mathbf{x}) := \sum_{k,l} x_kx_lw_{k,l} = \mathbf{x}^*W\mathbf{x}$. Let

$$\sigma^2(\mathbf{x}) = \sum_{k,l=1}^{K} x_kx_l\left(w_{k,l} - \omega(\mathbf{x})\right)^2;\ \overline{\sigma}^2\left(\mathbf{x}\right) = \sum_{k,l=1}^{K} x_kx_l\left(\frac{w_{k,l}}{\omega(\mathbf{x})} - 1\right)^2 = \frac{\sigma^2(\mathbf{x})}{\omega(\mathbf{x})^2} \tag{9}$$

be respectively the genotypic variance in absolute fitness and the diploid variance in relative fitness.

If we first define the frequency-dependent marginal fitness of A_k by $w_k(\mathbf{x}) = (W\mathbf{x})_k := \sum_l w_{k,l}x_l$, the marginal dynamics is given as in (3) by:

$$x'_k = x_k \frac{w_k(\mathbf{x})}{\omega(\mathbf{x})} = \frac{1}{\omega(\mathbf{x})} x_k (W\mathbf{x})_k =: p_k(\mathbf{x}),\ k = 1, ..., K. \tag{10}$$

In vector form (10) reads

$$\mathbf{x}' = \frac{1}{\omega(\mathbf{x})} D_{\mathbf{x}} W\mathbf{x} = \frac{1}{\omega(\mathbf{x})} D_{W\mathbf{x}}\mathbf{x} =: \mathbf{p}(\mathbf{x}),$$

where $\mathbf{p}$ maps S_K into S_K. Iterating, the time-t frequency distribution $\mathbf{x}(t)$ is the $t-$times composition of $\mathbf{p}$ applied to some $\mathbf{x}(0)$.

In the diploid case, assuming fitnesses to be multiplicative, say with $w_{k,l} = w_k w_l$, then selection acts on the gametes rather than on the genotypes. Observing $\frac{w_k(\mathbf{x})}{\mathbf{x}^* W\mathbf{x}} = \frac{w_k}{\sum_l w_l x_l}$, the dynamics (10) boils down to (3). However, the mean fitness in this case is $\omega(\mathbf{x}) = \left(\sum_l w_l x_l\right)^2$ and not $w(\mathbf{x}) = \sum_l w_l x_l$ as in the haploid case.

2.2.3. *Increase of mean fitness.* Again, the mean fitness $\omega(\mathbf{x})$, as a Lyapunov function, increases as time passes by. We indeed have

$$\Delta\omega(\mathbf{x}) = \omega(\mathbf{x}') - \omega(\mathbf{x}) = \frac{1}{\omega(\mathbf{x})^2} \sum_{k,l} x_k w_k(\mathbf{x}) w_{k,l} x_l w_l(\mathbf{x}) - \sum_{k,l} x_k w_{k,l} x_l \geq 0,$$

and vanishes only when the process has reached equilibrium.

Its partial rate of increase due to frequency shifts only is $\delta\omega(\mathbf{x}) := \sum_k \Delta x_k w_k(\mathbf{x})$. It satisfies

$$\frac{\delta\omega(\mathbf{x})}{\omega(\mathbf{x})} = \sum_k x_k \left(\frac{w_k(\mathbf{x})}{\omega(\mathbf{x})} - 1\right)^2 = \sum_k \frac{(\Delta x_k)^2}{x_k} = \frac{1}{2}\overline{\sigma}_A^2(\mathbf{x}), \tag{11}$$

where $\overline{\sigma}_A^2(\mathbf{x})$ is the allelic variance in relative fitness

$$\overline{\sigma}_A^2(\mathbf{x}) := 2\sum_{k=1}^{K} x_k \left(\frac{w_k(\mathbf{x})}{\omega(\mathbf{x})} - 1\right)^2. \tag{12}$$

2.2.4. *An alternative representation of the allelic dynamics.* There is an alternative vectorial representation of the dynamics (10) emphasizing its gradient-like character. Define the matrix $G(\mathbf{x}) = D_{\mathbf{x}} - \mathbf{x}\mathbf{x}^*$. It is symmetric, positive semi-positive whose quadratic form vanishes only for the constants. $G(\mathbf{x})$ is partially invertible on the space orthogonal to the constants with left-inverse

$$G(\mathbf{x})^{-1} = \left(I - \frac{1}{K}J\right) D_{\mathbf{x}}^{-1}.$$

Note $G(\mathbf{x})\,G(\mathbf{x})^{-1}\boldsymbol{\delta} \neq \boldsymbol{\delta}$. Looking for a left-inverse in the weaker sense of the quadratic form, that is satisfying

$$\boldsymbol{\delta}^* G(\mathbf{x})^{-1} G(\mathbf{x})\,\boldsymbol{\delta} = \boldsymbol{\delta}^* I \boldsymbol{\delta}$$

for all $\boldsymbol{\delta}$ with $|\boldsymbol{\delta}| = 0$, every $G(\mathbf{x})^{-1} = \left(I - \frac{\lambda}{K} J\right) D_{\mathbf{x}}^{-1}$ would do the job for any $\lambda \in \mathbb{R}$. In particular $\lambda = 0$.

Introduce the quantity $V_W(\mathbf{x}) = \frac{1}{2}\log\omega(\mathbf{x})$. Then, (10) may be recast as the gradient-like dynamics:

$$\Delta\mathbf{x} = \frac{1}{\omega(\mathbf{x})} G(\mathbf{x}) W \mathbf{x} = G(\mathbf{x}) \nabla V_W(\mathbf{x}), \tag{13}$$

with $|\Delta\mathbf{x}| = \mathbf{1}^*\Delta\mathbf{x} = 0$ as a result of $\mathbf{1}^* G(\mathbf{x}) = \mathbf{0}^*$. Note

$$\nabla V_W(\mathbf{x})^*\,\Delta\mathbf{x} = \nabla V_W(\mathbf{x})^*\,G(\mathbf{x})\,\nabla V_W(\mathbf{x}) \geq 0.$$

Based on [9], [10], the dynamics (13) is of gradient-type with respect to the Shashahani-Svirezhev distance metric given by

$$d_G(\mathbf{x},\mathbf{x}') = \left(\Delta\mathbf{x}^* G(\mathbf{x})^{-1}\,\Delta\mathbf{x}\right)^{1/2} = \left(\sum_{k=1}^{K} x_k^{-1}\,(\Delta x_k)^2\right)^{1/2}.$$

Its trajectories are perpendicular to the level surfaces of V_W with respect to this metric. From (11) and (12), $d_G(\mathbf{x},\mathbf{x}')$, which is the length of $\Delta\mathbf{x}$, is also the square-root of half the allelic variance (the standard deviation) in relative fitness.

3. The mutation mapping

We now briefly describe the frequency distribution dynamics when mutation is the only driving source of motion.
Assume alleles mutate according to the scheme: $A_k \to A_l$ with probability $\mu_{k,l} \in [0,1]$ satisfying $\mu_{k,k} = 0$ and $0 < \sum_{l\neq k}\mu_{k,l} \leq 1$ for all k. Let $M := \left[\mu_{k,l}\right]$ be the mutation pattern matrix; we shall assume that the non-negative matrix M is irreducible. We first consider the deterministic diploid model involving mutations.

3.1. **Only mutations.**

Considering first an updating mechanism of the frequencies where only mutations operate, we get

$$x_k' = x_k + \sum_{l\neq k}\mu_{l,k}x_l - x_k\sum_{l\neq k}\mu_{k,l},\ k = 1,...,K. \tag{14}$$

In matrix form, with M^* the transpose of M

$$\mathbf{x}' = \mathbf{x} + M^*\mathbf{x} - D_{M\mathbf{1}}\mathbf{x} =: \mathbf{M}\mathbf{x} =: \mathbf{p}_M(\mathbf{x}), \tag{15}$$

and the update of the frequencies with mutations is given by the affine transformation

$$\mathbf{M} := I - D_{M\mathbf{1}} + M^*.$$

We have $\mathbf{M} \geq 0$ and $\mathbf{M} = M^*$ if and only if M is stochastic, $M\mathbf{1} = \mathbf{1}$. Also $\mathbf{1}^*\mathbf{M} = \mathbf{1}^*$ and then $\mathbf{M}$ maps S_K into S_K because if $\mathbf{1}^*\mathbf{x} = 1$, then $\mathbf{1}^*\mathbf{x}' = \mathbf{1}^*\mathbf{M}\mathbf{x} = (\mathbf{M}^*\mathbf{1})^*\,\mathbf{x} = \mathbf{1}^*\mathbf{x} = 1$. The matrix $\mathbf{M}^*$ is stochastic and irreducible and so, by Perron-Frobenius theorem, it has a unique strictly positive probability

left-eigenvector associated to the real dominant eigenvalue 1. Let $\mathbf{x}_{eq}^*$ be this line-vector, then $\mathbf{x}_{eq}^* = \mathbf{x}_{eq}^*\mathbf{M}^*$, or $\mathbf{x}_{eq} = \mathbf{M}\mathbf{x}_{eq}$. Under the irreducibility assumption on M, the frequencies dynamics involving only mutations has a unique polymorphic equilibrium fixed point $\mathbf{x}_{eq} > \mathbf{0}$. When $\mathbf{M}$ is primitive then $\lim_{t\to\infty} \mathbf{M}^t = \mathbf{x}_{eq}\mathbf{1}^*$. This shows that

$$\mathbf{x} := \mathbf{x}(t) = \mathbf{M}^t\mathbf{x}(0) \underset{t\to\infty}{\to} \mathbf{x}_{eq}\mathbf{1}^*\mathbf{x}(0) = \mathbf{x}_{eq},$$

regardless of the initial condition $\mathbf{x}(0)$ belonging to S_K.

Note finally that from (15):

$$\Delta\mathbf{x} = (\mathbf{M} - I)\,\mathbf{x} =: \nabla V_{\mathbf{M}}(\mathbf{x}), \tag{16}$$

where $V_{\mathbf{M}}(\mathbf{x}) = \frac{1}{2}\mathbf{x}^*\,(\mathbf{M} - I)\,\mathbf{x}$ is the quadratic mutation potential. The probability right-eigenvector $\mathbf{x}_{eq}$ of $\mathbf{M}$ uniquely solves $\nabla V_{\mathbf{M}}\,(\mathbf{x}) = 0$ with $V_{\mathbf{M}}(\mathbf{x}_{eq}) = 0$, maximal.

3.2. **Remarks and special cases.** (i) Reversible mutations: Let $\mathbf{x}_{eq}$ solve $\mathbf{x}_{eq}^* = \mathbf{x}_{eq}^*\mathbf{M}^*$. Define

$$\overleftarrow{\mathbf{M}^*} = D_{\mathbf{x}_{eq}}^{-1}\mathbf{M}D_{\mathbf{x}_{eq}}.$$

We have $\overleftarrow{\mathbf{M}^*}\mathbf{1} = D_{\mathbf{x}_{eq}}^{-1}\mathbf{M}\mathbf{x}_{eq} = \mathbf{1}$, so $\overleftarrow{\mathbf{M}^*}$ is the stochastic matrix of the time-reversed process at equilibrium with invariant measure $\mathbf{x}_{eq}^*$. If $\overleftarrow{\mathbf{M}^*} = \mathbf{M}^*$, then the mutation pattern is said to be time-reversible. In this case

$$\mu_{k,l} = \mu_{l,k}\frac{x_{eq,k}}{x_{eq,l}}.$$

(ii) If $M = M^*$, then $\mathbf{M} = \mathbf{M}^*$ and $\mathbf{M}$ is doubly stochastic. In that case, $\mathbf{x}_{eq} = \frac{1}{K}\cdot(1,...,1)^* =: \mathbf{x}_b$. A model with symmetric mutations by assuming for instance multiplicative mutations: $\mu_{k,l} = \mu_k\mu_l$. In this case, with $\boldsymbol{\mu}$ the column vector of the μ_ks,

$$\mathbf{M} = I + \boldsymbol{\mu}\boldsymbol{\mu}^* - |\boldsymbol{\mu}|\,D_{\boldsymbol{\mu}}.$$

Alternatively, assuming $\mu_{k,l} = \mu \in \left(0, \frac{1}{K-1}\right]$ for all $k \neq l$ leads to $\mathbf{M} = \mu J + (1 - K\mu)\,I$ which is also symmetric.

It is not necessary that $M = M^*$ in order to have $\mathbf{M}$ doubly stochastic. It suffices to impose $M\mathbf{1} = M^*\mathbf{1}$. In that case although $\mathbf{M} \neq \mathbf{M}^*$, the overall input-output mutation probabilities attached to any state coincide and the equilibrium state again matches with the barycenter $\mathbf{x}_b$ of S_K.

(iii) (Kingman house of cards, [4]). Assume the mutation probabilities only depend on the terminal state, that is: $\mu_{k,l} = \mu_l$ for all $k \neq l$, still with $\mu_{k,k} = 0$. Let $\boldsymbol{\mu}^* = (\mu_1, ..., \mu_K)$. Then, $M = \mathbf{1}\boldsymbol{\mu}^* - D_{\boldsymbol{\mu}}$, $M\mathbf{1} = |\boldsymbol{\mu}|\cdot\mathbf{1} - \boldsymbol{\mu}$ where $\min\mu_k < |\boldsymbol{\mu}| := \boldsymbol{\mu}^*\mathbf{1} < 1 + \max\mu_k$, $\mathbf{M} = \boldsymbol{\mu}\mathbf{1}^* + (1 - |\boldsymbol{\mu}|)\,I$ and

$$\mathbf{x}' = \mathbf{M}\mathbf{x} = \mathbf{x} + M^*\mathbf{x} - D_{\mathbf{x}}M\mathbf{1} = \boldsymbol{\mu} + (1 - |\boldsymbol{\mu}|)\,\mathbf{x}. \tag{17}$$

The equilibrium state is $\mathbf{x}_{eq} = \boldsymbol{\mu}/\,|\boldsymbol{\mu}|$. Note that $|\boldsymbol{\mu}| \leq 1 + \frac{1}{K-1}$. This model is reversible. In this model the coordinates are decoupled: $x_k' = \mu_k + (1 - |\boldsymbol{\mu}|)\,x_k$, depends only on x_k.

(iv) Assume the mutation probabilities only depend on the initial state, that is: $\mu_{k,l} = \mu_k$ for all $l \neq k$. Then

$$\mathbf{M} = I - KD_{\boldsymbol{\mu}} + \mathbf{1}\boldsymbol{\mu}^*.$$

This mutation model is also reversible and the equilibrium state is

$$x_{eq,k} = \frac{1/\mu_k}{\sum_l 1/\mu_l}.$$

(v) (Random walk). In this case, M is tri-diagonal with $\mu_{k,l} = u_{\pm}\delta_{k,k\pm 1}$ and $\mathbf{x}_{eq}$ is a truncated geometric distribution with common ratio $\rho = u_+/u_-$:

$$x_{eq,k} = \frac{\rho^{k-1}}{\sum_{l=1}^{K}\rho^{l-1}},\ k = 1, ..., K.$$

This model is reversible.

(vi) (Cyclic mutation pattern). Here, $\mu_{k,l} = \mu_k\delta_{l,k+1}$, $k = 1, ..., K-1$ and $\mu_{K,l} = \mu_K\delta_{l,1}$. This model is not reversible and

$$x_{eq,k} = \frac{1/\mu_k}{\sum_l 1/\mu_l}.$$

4. Combining fitness and mutations

Let us now consider the dynamics driven both by fitness and mutation. There are two ways to combine the fitness and mutation effects. One (fitness-first), which is classical, consists in applying first the fitness mapping and then let mutation act on the result. The other (mutation-first), which seems to be less popular, consists in reversing the order.

4.1. **Fitness-first dynamics.**

It is typically obtained by applying first the fitness operator and then the mutation one to give the 'fitness-first' dynamics [11]:

$$\mathbf{x}' = \frac{1}{\mathbf{x}^* W\mathbf{x}}\mathbf{M}D_{W\mathbf{x}}\mathbf{x} = \frac{1}{\mathbf{x}^* W\mathbf{x}}\mathbf{M}D_{\mathbf{x}}W\mathbf{x}, \tag{18}$$

defining a new nonlinear transformation. Alternatively, $\mathbf{x}' = \mathbf{p}(\mathbf{x})$ where $\mathbf{p}(\mathbf{x}) = \frac{1}{\mathbf{x}^* W\mathbf{x}}\mathbf{M}D_{\mathbf{x}}W\mathbf{x}$ is the new mapping from S_K to S_K to consider. Component-wise, this is also as required

$$x'_k = \frac{1}{\omega(\mathbf{x})}\left(x_k w_k(\mathbf{x}) + \sum_{l\neq k}\mu_{l,k}w_l(\mathbf{x})x_l - x_k w_k(\mathbf{x})\sum_{l\neq k}\mu_{k,l}\right),\ k = 1, ..., K. \tag{19}$$

We have: $\mathbf{p}(\mathbf{e}_k) = \left(\mu_{k,1}, ..., \mu_{k,k-1}, 1 - \sum_{l\neq k}\mu_{k,l}, \mu_{k,k+1}, ..., \mu_{k,K}\right)^* \in S_K$ and so the extremal states $\mathbf{e}_k$ are not invariant under $\mathbf{p}$ and from the fixed-point theorem, there exists some equilibrium state in S_K. Using the representation (13) and (15):

$$\Delta\mathbf{x} = (\mathbf{M} - I)\,\mathbf{x} + \mathbf{M}G(\mathbf{x})\nabla V_W(\mathbf{x}) = \nabla V_{\mathbf{M}}(\mathbf{x}) + \mathbf{M}G(\mathbf{x})\nabla V_W(\mathbf{x}). \tag{20}$$

This is not a gradient-like dynamics in general because there is a competition between the mutation and fitness potentials $V_{\mathbf{M}}$ and V_W.

When $\mathbf{M} = I$ (no mutation) (20) boils down into (13) and when $W = J$ (no selection), (20) boils down into (16). When both $\mathbf{M} = I$ (no mutation) and $W = J$ (no selection), $\Delta\mathbf{x} = 0$ with corresponding neutral $\mathbf{p}(\mathbf{x}) = \mathbf{x}$.

4.2. **Mutation-first.** Because W was assumed symmetric ($W = W^*$), there is another way to combine the mutation-selection effects. It is obtained by applying first the mutation operator and then the fitness operator to give the 'mutation-first' dynamics:

$$\mathbf{x}' =: \mathbf{p}(\mathbf{x}) = \frac{1}{\omega(\mathbf{Mx})} D_{W\mathbf{Mx}}\mathbf{Mx}, \tag{21}$$

where $\omega(\mathbf{Mx}) := (\mathbf{Mx})^* W\mathbf{Mx}$. We have: $\mathbf{1}^*\mathbf{p}(\mathbf{x}) = (W\mathbf{Mx})^*\mathbf{Mx}/\omega(\mathbf{Mx}) = 1$ if and only if $W = W^*$ and under this condition, this new $\mathbf{p}(\mathbf{x})$ again maps the $K-$simplex S_K onto itself. The dynamics of $\mathbf{y} := \mathbf{Mx}$ is

$$\mathbf{y}' = \frac{1}{\omega(\mathbf{y})}\mathbf{M}D_{W\mathbf{y}}\mathbf{y},$$

which is of the form (18) and $\mathbf{x} = \mathbf{M}^{-1}\mathbf{y}$ may be recovered as an output from $\mathbf{y}$ only if $\mathbf{M}$ is invertible.

Component-wise, each component $p_k(\mathbf{x})$ may be read from

$$x_k' = \frac{(W\mathbf{Mx})_k}{\omega(\mathbf{Mx})}\left(x_k\left(1 - \sum_{l\neq k}\mu_{k,l}\right) + \sum_{l\neq k}\mu_{l,k}x_l\right),\ k = 1, ..., K. \tag{22}$$

Equation (21) may also be recast as:

$$\Delta\mathbf{x} = (\mathbf{M} - I)\,\mathbf{x} + \frac{1}{\omega(\mathbf{Mx})}G(\mathbf{Mx})W\mathbf{Mx} = \nabla V_{\mathbf{M}}(\mathbf{x}) + G(\mathbf{Mx})\nabla V_W(\mathbf{Mx}), \tag{23}$$

where, as before, $V_W(\mathbf{x}) = \log\omega(\mathbf{x})/2$. The mean fitness function appearing in (23) is

$$\omega(\mathbf{Mx}) := (\mathbf{Mx})^* W\mathbf{Mx} = \mathbf{x}^*\mathbf{M}^* W\mathbf{Mx},$$

or else, the new fitness matrix to consider is $W_{\mathbf{M}} := \mathbf{M}^* W\mathbf{M}$, which is itself symmetric because $W = W^*$. Just like (20), (23) neither is a gradient-like dynamics.

Although natural, this alternative 'mutation-first' way to combine mutations and fitness effects seems to have been less studied in the literature.

5. Multiplicative fitness

We now focus on the multiplicative fitness model. Following the previous observations, we shall distinguish two cases.

5.1. **Fitness-first.** In the haploid case or in the diploid case when fitnesses are multiplicative, $w_{k,l} = w_k w_l$, with $w(\mathbf{x}) = \sum_l w_l x_l$

$$w(\mathbf{x})\, x'_k = x_k w_k + \sum_{l \neq k} \mu_{l,k} w_l x_l - x_k w_k \sum_{l \neq k} \mu_{k,l},\ k = 1, ..., K,$$

or

$$\mathbf{x}' = \frac{1}{w(\mathbf{x})} \mathbf{M} D_{\mathbf{x}} \mathbf{w} = \frac{1}{w(\mathbf{x})} \mathbf{M} D_{\mathbf{w}} \mathbf{x}, \tag{24}$$

where $\mathbf{w}$ is the constant column-vector of the w_ks and $w(\mathbf{x}) = \mathbf{w}^* \mathbf{x}$.

When dealing with multiplicative fitnesses models, we shall assume $\min_k w_k > 0$, $\max_k w_k = 1$ and the second largest $w_k < 1$.

The image of the extremal states $\mathbf{x} = \mathbf{e}_l$ by the transformation $\mathbf{x} \to \mathbf{x}'$ reduces to $\mathbf{M}\mathbf{e}_l$ which belongs to the interior of S_K. In that case, there exists a unique, globally stable polymorphic equilibrium state which is the fixed-point of (24). This follows from the Perron-Frobenius theorem commented in the forthcoming paragraph. Recall that in the absence of mutations, the multiplicative fitness model cannot have a polymorphic equilibrium state.

5.1.1. *Polymorphic equilibrium and steady mean fitness.* Let $A = D_{\mathbf{w}} \mathbf{M}^*$, the latter 'selection-first' recurrence may be recast as

$$\mathbf{x}^*(t+1) = \frac{1}{\mathbf{x}^*(t) A \mathbf{1}} \mathbf{x}^*(t) A. \tag{25}$$

Under our assumptions on $\mathbf{w}$, $A\mathbf{1} = \mathbf{w}$ and therefore A is sub-stochastic. By iteration

$$\mathbf{x}^*(t) = \frac{1}{\mathbf{x}^*(0) A^t \mathbf{1}} \mathbf{x}^*(0) A^t.$$

When $\mathbf{M}$ is primitive, so is $A \geq 0$ which has Perron-Frobenius left and right probability eigenvectors $\mathbf{x}^*_A > 0$ and $\mathbf{y}_A > 0$ associated to its largest eigenvalue $1 > \rho_A > 0$.

Then, $\lim_{t \to \infty} \left(\frac{1}{\rho_A} A \right)^t = \frac{1}{\sum_k x_{A,k} y_{A,k}} \mathbf{y}_A \mathbf{x}^*_A$ showing that,

$$\forall \mathbf{x}(0) : \ \lim_{t \to \infty} \mathbf{x}(t) = \mathbf{x}_A > \mathbf{0},$$

which is the required limiting polymorphic state. The value of ρ_A (respectively ρ_A^2) is the limiting haploid (diploid) mean fitness because: $w_A := w(\mathbf{x}_A) = \mathbf{w}^* \mathbf{x}_A = \mathbf{x}^*_A A \mathbf{1} = \rho_A$. When looking at the equivalent reformulation (25) of (24), $\mathbf{x}_A$ can be interpreted as a quasi-stationary distribution as developed now. For the precise definition, see [12].

5.1.2. *A stochastic interpretation of the deterministic dynamics (25).* A vector $\mathbf{x}$ of S_K can be thought of as a probability vector. The dynamical equation (24), as a nonlinear update mapping from S_K to S_K, may be viewed as the discrete-time nonlinear master equation of some Markov process whose construction we now give. We shall need to introduce an extra state, say $\partial = \{0\}$ which will be absorbing for

the process we shall now construct. It will be useful to extend the matrix A to ∂ in the following way:

$$A_{k,0} = 1 - \sum_{l=1}^{K} A_{k,l},\ A_{0,l} = \delta_{l,0}.$$

Let then $L(t)$ be the random labels distribution of an individual at time t, with enlarged state-space $\{0, 1, ..., K\}$. Let $(U_t : t = 1, 2)$ be an i.i.d. driving sequence of uniformly distributed random variables on $[0, 1]$. Consider the random evolution equation

$$\mathbf{1}\left(L(t+1) = l', L(t) = l\right) = \mathbf{1}\left(U_{t+1} < A_{L(t),l'}\right) \cdot \mathbf{1}(L(t) = l),$$

where $l, l' \in \{0, 1, ..., K\}$.

From this construction $(L(s) : s \leq t)$ is measurable with respect to $(U_s : s \leq t)$ and we get

$$\mathbb{P}\left(L(t+1) = l' \mid L(t)\right) = A_{L(t),l'}. \tag{26}$$

Let τ be the first time that $L(t)$ hits the absorbing state $\partial = \{0\}$. Using the extinction time (26) may be recast as

$$\mathbb{P}\left(L(t+1) = k, \tau > t+1 \mid L(t)\right) = A_{L(t),k}\mathbf{1}(\tau > t).$$

Putting $z_k(t) := \mathbb{P}\left(L(t) = k, \tau > t\right)$, we get an un-normalized version of (25):

$$z_k\left(t+1\right) = \mathbb{E}\left(A_{L(t),k}\mathbf{1}(\tau > t)\right) = \left(\mathbf{z}^*(t)A\right)_k, k \in \{1, ..., K\}.$$

We clearly have

$$\lim_{t\to\infty} \rho_A^{-t}\mathbb{P}_k(\tau > t) = \frac{y_{A,k}}{\sum_k x_{A,k}y_{A,k}}, \tag{27}$$

and so $1 > \mathbb{P}\left(\tau > t\right) \sim \rho_A^t \to 0$ geometrically fast. From the last expression, the right-hand-side may be interpreted as the propensity of a type-k allele to survive to its fate: the eventual extinction. If $y_{A,k} > y_{A,l}$ indeed, the extinction time of the process started at k is larger than the one started at l (has larger survival asymptotic tails). We shall call $\mathbf{y}_A$ the survival probability vector.

Defining the normalized conditional probabilities

$$x_k\left(t\right) = \frac{z_k(t)}{\sum_{k=1}^{K} z_k(t)} = \mathbb{P}\left(L(t) = k \mid \tau > t\right),$$

we obtain the normalized haploid dynamics (25)

$$x_k' = \frac{\left(\mathbf{x}^*A\right)_k}{\sum_{k=1}^{K}\left(\mathbf{x}^*A\right)_k},\ k \in \{1, ..., K\}.$$

It may now be viewed as the nonlinear master equation of some stochastic Markovian process. In view of this construction, the vector $\mathbf{x}_A$ is the quasi-stationary distribution of $L(t)$ given $\tau > t$.

We note that the appeal to the coffin state ∂ was a necessary step to understand the normalization $z_k \to x_k$, and the stochastic interpretation of (25) allows to give sense to the right eigenvector $\mathbf{y}_A$ of A.

Clearly the above construction can be done for N particles, in particular $N = 2$ in the diploid case.

5.2. **Mutation-first.** If instead of (24) the dynamics is of the type 'mutation-first'

$$\mathbf{x}' = \frac{1}{\mathbf{w}^* \mathbf{M}\mathbf{x}} D_{\mathbf{w}} \mathbf{M}\mathbf{x}, \tag{28}$$

because the mutation operator was applied first in the composition of the fitnesses and mutation effects, the latter recurrence may be recast as:

$$\mathbf{x}^*(t+1) = \frac{1}{\mathbf{x}^*(t) B\mathbf{1}} \mathbf{x}^*(t) B, \tag{29}$$

where now $B = \mathbf{M}^* D_{\mathbf{w}}$. By iteration $\mathbf{x}^*(t) = \frac{1}{\mathbf{x}(0)^* B^t \mathbf{1}} \mathbf{x}^*(0) B^t$.

5.2.1. *Equilibrium and mean fitness at equilibrium.* Let $\mathbf{x}_B > \mathbf{0}$ and $\mathbf{y}_B > \mathbf{0}$, be now the left and right probability Perron-Frobenius eigenvector of B associated to its largest eigenvalue $\rho_B > 0$. We have $\lim_{t\to\infty} \left(\frac{1}{\rho_B} B\right)^t = \frac{1}{\sum_k x_{B,k} y_{B,k}} \mathbf{y}_B \mathbf{x}_B^*$ showing that

$$\forall \mathbf{x}(0) : \lim_{t\to\infty} \mathbf{x}(t) = \mathbf{x}_B > \mathbf{0}.$$

We have $B = D_{\mathbf{w}}^{-1} A D_{\mathbf{w}}$ and so B is diagonally similar to A. Therefore $\rho_B = \rho_A$ and $\mathbf{x}_B = D_{\mathbf{w}} \mathbf{x}_A / w_A$ together with $\mathbf{y}_B = D_{\mathbf{w}}^{-1} \mathbf{y}_A / \left| D_{\mathbf{w}}^{-1} \mathbf{y}_A \right|$.

The limiting equilibrium mean fitness is now $w_B := \mathbf{w}^* \mathbf{x}_B = \mathbf{x}_B^* \mathbf{w}$ (or w_B^2 in the diploid case). Recalling $\mathbf{x}_B = D_{\mathbf{w}} \mathbf{x}_A / w_A$, we have

$$w_B = \frac{\sum_k w_k^2 x_{A,k}}{\sum_k w_k x_{A,k}}.$$

Since $w_A = \sum_k w_k x_{A,k} = \rho_A$, under the multiplicative fitness hypothesis we obtain:

Proposition 1. $\rho_A = w_A < w_B < 1$.

Remark: The quantity $\overline{\sigma}^2(\mathbf{x}_A) = \sum_k \left(\frac{w_k}{w_A} - 1\right)^2 x_{A,k}$ is the variance in relative fitness at equilibrium for the model A. We therefore have

$$\overline{\sigma}^2(\mathbf{x}_A) = \frac{w_B}{w_A} - 1 \text{ or } w_B = w_A \left(1 + \overline{\sigma}^2(\mathbf{x}_A)\right). \blacklozenge$$

The equilibrium fitness of the second model is larger than the one of the first. Without mutations, only the fittest state, say $\{K\}$ under our hypotheses, will survive, leading to an equilibrium mean fitness equal to $(0, 0, ..., 0, 1)\, \mathbf{w}^* = w_K = 1$. Therefore, both mutation models lead to a decrease of the equilibrium mean fitness, when compared to the one without mutations. However, the first model involves mutations which are more deleterious than the ones relative to the second one where mutations appear more advantageously.

Note finally that $\mathbf{w}^* \mathbf{M}\mathbf{x}_B = \mathbf{w}^* D_{\mathbf{w}}^{-1} B^* \mathbf{x}_B = \rho_B \mathbf{w}^* D_{\mathbf{w}}^{-1} \mathbf{x}_B = \rho_B \mathbf{1}^* \mathbf{x}_B = \rho_B = \rho_A$ and, since $B = \mathbf{M}^* D_{\mathbf{w}}$ then $B D_{\mathbf{w}}^{-1} = \mathbf{M}^*$ so that $B\mathbf{w}^{-1} = \mathbf{1}$ if $\mathbf{w}^{-1} := \left(w_1^{-1}, ..., w_K^{-1}\right)^*$ is the reciprocal fitness vector. As a result, $\mathbf{x}_B^* B \mathbf{w}^{-1} = \rho_B \mathbf{x}_B^* \mathbf{w}^{-1} = 1$ so that $\mathbf{x}_B^* \mathbf{w}^{-1} := \sum_k \frac{x_{B,k}}{w_k} = \rho_B^{-1}$.

5.2.2. *A stochastic interpretation of the deterministic dynamics (29).* We can repeat the above construction substituting B for A and we are done.

5.2.3. *The stochastic dominations* $\mathbf{x}_B \succeq_{st} \mathbf{x}_A$ *and* $\mathbf{y}_A \succeq_{st} \mathbf{y}_B$. For two $K-$dimensional probability vectors $\mathbf{a}$ and $\mathbf{b}$, we put $\mathbf{a} \succeq_{st} \mathbf{b}$ if for each l

$$\sum_{k=1}^{l} a_k \leq \sum_{k=1}^{l} b_k.$$

Proposition 2. *We have* $\mathbf{x}_B \succeq_{st} \mathbf{x}_A$.

Proof: $\mathbf{x}_B = D_{\mathbf{w}}\mathbf{x}_A/w_A$ and therefore

$$x_{B,k} = \frac{w_k x_{A,k}}{\sum_{k=1}^{K} w_k x_{A,k}}.$$

With $\alpha_l := \sum_{k=1}^{l} x_{A,k}$, we have

$$\sum_{k=1}^{K} w_k x_{A,k} = \alpha_l \sum_{k=1}^{l} w_k \frac{x_{A,k}}{\sum_{k=1}^{l} x_{A,k}} + (1-\alpha_l) \sum_{k=l+1}^{K} w_k \frac{x_{A,k}}{\sum_{k=l+1}^{K} x_{A,k}}.$$

Since $\sum_{k=1}^{l} w_k \frac{x_{A,k}}{\sum_{k=1}^{l} x_{A,k}} \in (w_1, w_l)$ and $\sum_{k=l+1}^{K} w_k \frac{x_{A,k}}{\sum_{k=l+1}^{K} x_{A,k}} \in (w_{l+1}, w_K)$,

$$\sum_{k=1}^{K} w_k x_{A,k} \geq \sum_{k=1}^{l} w_k \frac{x_{A,k}}{\sum_{k=1}^{l} x_{A,k}},$$

and therefore

$$\sum_{k=1}^{l} x_{A,k} \geq \frac{\sum_{k=1}^{l} w_k x_{A,k}}{\sum_{k=1}^{K} w_k x_{A,k}} = \sum_{k=1}^{l} x_{B,k},$$

which means $\mathbf{x}_B \succeq_{st} \mathbf{x}_A$. ♦

We point out that we used the order $0 < w_1 \leq ... \leq w_K = 1$ on $\mathbf{w}$. Would we have considered the reverse order, we would get the opposite domination relationship.

We also have the following stochastic domination property between the two survival probability vectors:

Corollary 3. $\mathbf{y}_A \succeq_{st} \mathbf{y}_B$.

Proof: Because B and A are diagonally similar, we also have $\mathbf{y}_A = D_{\mathbf{w}}\mathbf{y}_B/\left|D_{\mathbf{w}}\mathbf{y}_B\right|$ and the same argument applies substituting $(\mathbf{y}_A, \mathbf{y}_B)$ for $(\mathbf{x}_B, \mathbf{x}_A)$ in the previous proof. ♦

5.3. **Symmetric mutations.** When $\mathbf{M} = \mathbf{M}^*$, mutations are symmetric and $B = A^*$. Therefore $\mathbf{x}_B = \mathbf{y}_A$ and $\mathbf{y}_B = \mathbf{x}_A$. The left (right) probability eigenvector of B matches with the right (left) probability eigenvector of A. In this case, there is a stochastic domination property between the left and right eigenvectors of both models, namely

Proposition 4. *If mutations are symmetric* $\mathbf{y}_A \succeq_{st} \mathbf{x}_A$ *and* $\mathbf{x}_B \succeq_{st} \mathbf{y}_B$.

5.4. **Reversible mutations.** When dealing with reversible mutations with equilibrium distribution $\mathbf{x}_{eq}$, we show now that the right eigenvector of A (or B) can be computed from the left eigenvector by using an appropriate Schur product involving $\mathbf{w}$ and $\mathbf{x}_{eq}$.

Let A be an irreducible non-negative matrix. Let $(\mathbf{x}_A, \mathbf{y}_A)$ be the left and right probability eigenvectors of A, associated to the spectral radius ρ_A of A.

If there exists a positive vector $\boldsymbol{\eta}$ such that $\eta_k A_{k,l} = \eta_l A_{l,k}$ for all k, l, A is said to be reversible with respect to $\boldsymbol{\eta}$.

Consider the stochastic matrix

$$\widetilde{A} = \rho_A^{-1} D_{\mathbf{y}_A}^{-1} A D_{\mathbf{y}_A}. \tag{30}$$

Its left probability invariant measure is easily seen to be $\mathbf{x}_{\widetilde{A}} = D_{\mathbf{y}_A}\mathbf{x}_A / \sum_k x_{A,k} y_{A,k}$, which is the normalized Schur product of $\mathbf{x}_A$ and $\mathbf{y}_A$. We have

Lemma 5. *If A is reversible with respect to $\boldsymbol{\eta}$, then $\widetilde{A}$ is reversible with respect to $D^2_{\mathbf{y}_A}\boldsymbol{\eta}$, the Schur product of $\mathbf{y}_A$, $\mathbf{y}_A$ and $\boldsymbol{\eta}$.*

Proof:

$$\widetilde{A}_{l,k} = \rho_A^{-1} A_{l,k} \frac{y_{A,k}}{y_{A,l}} = \rho_A^{-1} \frac{\eta_k}{\eta_l} A_{k,l} \frac{y_{A,k}}{y_{A,l}} = \frac{\eta_k}{\eta_l} \widetilde{A}_{k,l} \frac{y^2_{A,k}}{y^2_{A,l}}. \blacklozenge$$

As an illustration, we shall consider the fitness-first dynamics for which $A = D_{\mathbf{w}}\mathbf{M}^*$. This A will be reversible with respect to $\boldsymbol{\eta}$ if and only if $\mathbf{M}^*$ itself is reversible. Indeed, $\boldsymbol{\eta}$ must satisfy

$$\eta_k w_k \mu_{k,l} = \eta_l w_l \mu_{l,k}, \tag{31}$$

and if this is the case $\mathbf{M}^*$ must be reversible with respect to $\mathbf{x}_{eq}$, leading to $\eta_k = x_{eq,k}/w_k$.

In case $\mathbf{M}^*$ is reversible, $\widetilde{A}$ is reversible with respect to $D^2_{\mathbf{y}_A}\boldsymbol{\eta}$ with entries proportional to $\left(x_{eq,k} y^2_{A,k}\right)/w_k$. But this must be the invariant measure of $\widetilde{A}$ which, up to a normalizing constant, is $D_{\mathbf{y}_A}\mathbf{x}_A$ with un-normalized entries $x_{A,k}y_{A,k}$. For this point, see also [13]. We conclude:

Proposition 6. *If the mutation matrix is reversible with respect to $\mathbf{x}_{eq}$, then $\widetilde{A}$ is reversible with respect to $D^2_{\mathbf{y}_A} D_w^{-1}\mathbf{x}_{eq} = D_w^{-1} D^2_{\mathbf{y}_A}\mathbf{x}_{eq}$ and*

$$\mathbf{x}_A = \frac{D_w^{-1} D_{\mathbf{x}_{eq}} \mathbf{y}_A}{\left| D_w^{-1} D_{\mathbf{x}_{eq}} \mathbf{y}_A \right|}.$$

Example: If $\mathbf{M}^*$ is symmetric, it is reversible with respect to the uniform measure $x_{eq,k} = 1/K$. Therefore $\mathbf{x}_A = D_w^{-1}\mathbf{y}_A / \left| D_w^{-1}\mathbf{y}_A \right|$. Up to a normalization constant, we have

$$\left(x_{eq,k} y^2_{A,k}\right)/w_k = y^2_{A,k}/w_k = w_k x^2_{A,k} = x_{A,k} y_{A,k},$$

the Schur product of $\mathbf{x}_A$ and $\mathbf{y}_A$.

Substituting B for A and $(\mathbf{x}_B, \mathbf{y}_B)$ for $(\mathbf{x}_A, \mathbf{y}_A)$, the same holds true for $B = \mathbf{M}^* D_{\mathbf{w}}$ and we get

Corollary 7. *If the mutation matrix is reversible with respect to $\mathbf{x}_{eq}$, then $\widetilde{B} = \rho_B^{-1} D_{\mathbf{y}_B}^{-1} B D_{\mathbf{y}_B}$ is reversible with respect to $D_{\mathbf{y}_B}^2 D_w \mathbf{x}_{eq} = D_w D_{\mathbf{y}_B}^2 \mathbf{x}_{eq}$ and*

$$\mathbf{x}_B = \frac{D_w D_{\mathbf{x}_{eq}} \mathbf{y}_B}{\left| D_w D_{\mathbf{x}_{eq}} \mathbf{y}_B \right|} \text{ or } x_{B,k} = \frac{w_k x_{eq,k} y_{B,k}}{\sum_k w_k x_{eq,k} y_{B,k}}.$$

6. Multiplicative fitness and the House of Cards condition

We shall again distinguish two cases.

6.1. **Fitness-first.** Assume the house of cards condition holds, leading to: $A = D_{\mathbf{w}} \mathbf{M}^*$, with $\mathbf{M}^* = \mathbf{1}\boldsymbol{\mu}^* + (1 - |\boldsymbol{\mu}|) I$. In this case, the computations become more explicit. Since for all k, $\sum_{l \neq k} \mu_{k,l} = \sum_{l \neq k} \mu_l = |\boldsymbol{\mu}| - \mu_k \leq 1$, we have

$$|\boldsymbol{\mu}| \leq \frac{K}{K-1} \text{ and } 1 - |\boldsymbol{\mu}| \geq -\frac{1}{K-1}.$$

Under the multiplicative fitness and the house of cards conditions, Equation (19) reads

$$x'_k = \frac{1}{w(\mathbf{x})} \left(x_k w_k + \mu_k \sum_{l \neq k} w_l x_l - x_k w_k \sum_{l \neq k} \mu_l \right) = \mu_k + \frac{1 - |\boldsymbol{\mu}|}{\mathbf{x}^* \mathbf{w}} w_k x_k.$$

The equilibrium frequency distribution is therefore the solution to the equation

$$\mathbf{x}_A = \frac{1 - |\boldsymbol{\mu}|}{\mathbf{x}_A^* \mathbf{w}} D_{\mathbf{w}} \mathbf{x}_A + \boldsymbol{\mu},$$

which, since $\rho_A = \mathbf{x}_A^* \mathbf{w}$, is exactly seen to be:

$$x_{A,k} = \frac{\mu_k}{1 - (1 - |\boldsymbol{\mu}|) \, w_k / \rho_A}, \; k = 1, ..., K, \tag{32}$$

where $\rho_A \in (0, 1)$ is such that $\sum_k x_{A,k} = 1$. See [4] where these results appear first. Alleles A_k with largest frequencies are those for which both (μ_k, w_k) are large. The equilibrium mean fitness is

$$\mathbf{x}_A^* \mathbf{w} = \sum_k \frac{\mu_k w_k}{1 - (1 - |\boldsymbol{\mu}|) \, w_k / \rho_A} = \rho_A, \tag{33}$$

the spectral radius of A and $A = D_{\mathbf{w}} \mathbf{M}^*$ satisfying $A\mathbf{1} = \mathbf{w}$. Because $\mathbf{x}_A > \mathbf{0}$, we have:

$$w_A = \rho_A > (1 - |\boldsymbol{\mu}|) \max_k w_k. \tag{34}$$

If $|\boldsymbol{\mu}| = 1$, $x_{A,k} = \mu_k$, $k = 1, ..., K$ and $\rho_A = \sum_k \mu_k w_k$.

6.2. **Mutation-first.** When fitnesses are multiplicative: $W = \mathbf{w}\mathbf{w}^*$ is symmetric, (21) is also $\mathbf{x}' =: \mathbf{p}(\mathbf{x}) = \frac{1}{\mathbf{w}^*\mathbf{M}\mathbf{x}} D_{\mathbf{w}}\mathbf{M}\mathbf{x}$. With $\mathbf{w} = (w_1, ..., w_k)^*$, this simplifies to give:

$$x'_k = \frac{w_k}{\mathbf{w}^*\mathbf{M}\mathbf{x}} \left(x_k \left(1 - \sum_{l \neq k} \mu_{k,l} \right) + \sum_{l \neq k} \mu_{l,k} x_l \right), \; k = 1, ..., K. \tag{35}$$

When the house of cards condition holds, $B = \mathbf{M}^* D_{\mathbf{w}}$, with $\mathbf{M}^* = \mathbf{1}\boldsymbol{\mu}^* + (1 - |\boldsymbol{\mu}|) I$. Equation (35) further simplifies to:

$$x'_k = \frac{w_k}{\mathbf{w}^*\boldsymbol{\mu} + (1 - |\boldsymbol{\mu}|)\, \mathbf{w}^*\mathbf{x}} \left(\mu_k + x_k \left(1 - |\boldsymbol{\mu}| \right) \right), \; k = 1, ..., K. \tag{36}$$

From Equation (36), the equilibrium frequency distribution is the solution to the equations

$$x_{B,k} = \frac{w_k}{\mathbf{w}^*\boldsymbol{\mu} + (1 - |\boldsymbol{\mu}|)\, \mathbf{x}_B^*\mathbf{w}} \left(\mu_k + x_{B,k} \left(1 - |\boldsymbol{\mu}| \right) \right), \; k = 1, ..., K,$$

which is exactly seen to be:

$$x_{B,k} = \frac{w_k \mu_k}{\mathbf{w}^*\boldsymbol{\mu} + (1 - |\boldsymbol{\mu}|)\, (w_B - w_k)}, \; k = 1, ..., K, \tag{37}$$

where $w_B = \mathbf{x}_B^*\mathbf{w} \in (0, 1)$ is such that $\sum_k x_{B,k} = 1$.

Alleles A_k with largest frequencies are those for which the product $w_k\mu_k$ is largest. Because A and B are diagonally similar, we have $\mathbf{x}_B = D_{\mathbf{w}}\mathbf{x}_A / w_A$ where $w_A = \rho_A = \rho_B$. From the expression (32) of $x_{A,k}$, we get the alternative expression

$$x_{B,k} = \frac{w_k \mu_k}{\rho_B - (1 - |\boldsymbol{\mu}|)\, w_k}, \; k = 1, ..., K. \tag{38}$$

We also have

$$\mathbf{x}_B^*\mathbf{w} = \sum_k \frac{w_k^2 \mu_k}{\rho_B - (1 - |\boldsymbol{\mu}|)\, w_k} = w_B, \tag{39}$$

the equilibrium mean fitness under $B = \mathbf{M}^* D_{\mathbf{w}}$.

Comparing the two expressions (37) and (38) of $x_{B,k}$, this suggests that $\rho_B = \mathbf{w}^*\boldsymbol{\mu} + (1 - |\boldsymbol{\mu}|)\, w_B$. Thus $w_A = \rho_A = \rho_B = \mathbf{w}^*\boldsymbol{\mu} + (1 - |\boldsymbol{\mu}|)\, w_B$ and

$$w_B - w_A = |\boldsymbol{\mu}|\, w_B - \mathbf{w}^*\boldsymbol{\mu} > 0. \tag{40}$$

As a result

$$w_B > \frac{1}{|\boldsymbol{\mu}|} \mathbf{w}^*\boldsymbol{\mu} \tag{41}$$

gives a lower bound for w_B in terms of the average of $\mathbf{w}$ with respect to the mutation equilibrium probability measure $\mathbf{x}_{eq} = \boldsymbol{\mu} / |\boldsymbol{\mu}|$.

Because $\mathbf{x}_B > \mathbf{0}$, from (37), we also have:

$$((1 - |\boldsymbol{\mu}|)\, w_k - \mathbf{w}^*\boldsymbol{\mu}) < (1 - |\boldsymbol{\mu}|)\, w_B, \; \forall k.$$

If $1 - |\boldsymbol{\mu}| > 0$, this means $w_B > \max_k w_k - \mathbf{w}^*\boldsymbol{\mu} / (1 - |\boldsymbol{\mu}|)$ whereas if $1 - |\boldsymbol{\mu}| < 0$ this means $w_B < \min_k w_k + \mathbf{w}^*\boldsymbol{\mu} / (|\boldsymbol{\mu}| - 1)$. When $|\boldsymbol{\mu}| = 1$, $w_B = \sum_k w_k^2 \mu_k / \sum_k w_k \mu_k$ is explicit, together with

$$x_{B,k} = \frac{w_k \mu_k}{\sum_k w_k \mu_k}, k = 1, ..., K.$$

Recalling that when $|\boldsymbol{\mu}| = 1 : w_A = \rho_A = \sum_k \mu_k w_k$, we can check in this particular case that: $w_B > w_A$.

7. Alternative conditional models for the evolutionary dynamics

The interpretation of (24) and of (29) in terms of a stochastic process conditioned on not being currently absorbed in ∂ suggests that other conditional models for the evolutionary dynamics involving multiplicative fitness and mutations could also be worth investigating.

Consider first the fitness-first model (24) driven by $A = D_{\mathbf{w}}\mathbf{M}^*$. Let $\widehat{A} = D_{A\mathbf{1}}^{-1}A$ and consider the updating dynamics on the simplex

$$\mathbf{x}^*(t+1) = \mathbf{x}^*(t)\,\widehat{A}. \tag{42}$$

Because $A\mathbf{1} = \mathbf{w}$, we have $\widehat{A} = \mathbf{M}^*$ which is the pure mutation stochastic matrix. Using the terms of the stochastic interpretation of (24), we have: $\widehat{A}_{k,l} = \mathbb{P}_{L(0)=k}(L(1) = l \mid \tau > 1)$ which is the transition matrix of a one-step conditioned process. Therefore conditioning locally $L(t)$ on non-extinction brings one back to the pure underlying mutation model with $\mathbf{x}_{\widehat{A}} = \mathbf{x}_{eq}$.

Let us consider a similar conditioning but for the mutation-first dynamics. With now $B = \mathbf{M}^* D_{\mathbf{w}}$, let $\widehat{B} = D_{B\mathbf{1}}^{-1}B$ and consider the dynamics

$$\mathbf{x}^*(t+1) = \mathbf{x}^*(t)\,\widehat{B}. \tag{43}$$

We have $\widehat{B} = D_{B\mathbf{1}}^{-1}B = D_{\mathbf{M}^*\mathbf{w}}^{-1}\mathbf{M}^* D_{\mathbf{w}}$ which is stochastic but cannot be reduced to mutation effects in general. This is an additional illustration of the differences between the two models based on A or B. Note that $\widehat{B}$ and $\widehat{A}$ are not diagonally similar.

Let now $\widetilde{A} := \rho_A^{-1} D_{\mathbf{y}_A}^{-1} A D_{\mathbf{y}_A}$ and consider the dynamics on the simplex

$$\mathbf{x}^*(t+1) = \mathbf{x}^*(t)\widetilde{A}. \tag{44}$$

We have: $\widetilde{A}_{k,l} = \lim_{s\to\infty} \mathbb{P}_{L(0)=k}(L(1) = l \mid \tau > s)$ which is the transition matrix of a process conditioned on not getting extinct in the remote future (see [14]). Therefore conditioning globally $L(t)$ on non-extinction in the far future brings one back to a standard (linear) Chapman-Kolmogorov evolution equation. This conditioning being more stringent than the one involved in (24), one expects its limiting frequency distribution (which is $\mathbf{x}_{\widetilde{A}} = D_{\mathbf{x}_A}\mathbf{y}_A / \sum_k x_{A,k} y_{A,k}$) to stay away more significantly from the origin $\{0\}$. For this model, the mean fitness at equilibrium will be:

$$w_{\widetilde{A}} = \mathbf{w}^* D_{\mathbf{x}_A}\mathbf{y}_A = \frac{\sum_k w_k x_{A,k} y_{A,k}}{\sum_k x_{A,k} y_{A,k}}.$$

When mutations are reversible,

$$x_{\widetilde{A},k} = \frac{w_k x_{eq,k}^{-1} x_{A,k}^2}{\sum_k w_k x_{eq,k}^{-1} x_{A,k}^2}.$$

Similar conclusions can be drawn if we define $\widetilde{B} := \rho_B^{-1} D_{\mathbf{y}_B}^{-1} A D_{\mathbf{y}_B}$.

The main interest is that in both conditioning (either local or global), the deterministic updating mechanisms are now linear in sharp contrast with (24) and (29) involving rational updating mechanisms.

Acknowledgments: The authors are indebted for support of the Basal Conicyt project and S.M. to the Guggenheim fellowship.

References

[1] Ewens, W. J. "Mathematical population genetics. I. Theoretical introduction," Second edition. Interdisciplinary Applied Mathematics, 27. Springer-Verlag, New York, 2004.

[2] Bürger, R. "The mathematical theory of selection, recombination, and mutation," Wiley Series in Mathematical and Computational Biology. John Wiley & Sons, Ltd., Chichester, 2000.

[3] Karlin, S. "Mathematical models, problems, and controversies of evolutionary theory," Bull. Amer. Math. Soc. (N.S.) Volume 10, No 2, 1984, pp. 221-273.

[4] Kingman, J. F. C. "Mathematics of genetic diversity," CBMS-NSF Regional Conference Series in Applied Mathematics, 34. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa., 1980. vii+70 pp. ISBN: 0-89871-166-5.

[5] Hermisson, J., Redner, O., Wagner, H. and Baake E. "Mutation-Selection Balance: Ancestry, Load and Maximum Principle," Theoretical Population Biology, Volume 62, Issue 1, 2002, pp. 9-46.

[6] Baake E. and Georgii, H-O. "Mutation, selection, and ancestry in branching models: a variational approach," J. Math. Biol. 54, 2007, pp. 257-303.

[7] Sella, G. and Hirsh, A. E. "The application of statistical physics to evolutionary biology," PNAS 102(27), 2005, pp. 9541-9546.

[8] Champagnat, N., Ferrière, R. and Méleard, S. "From individual stochastic processes to macroscopic models in adaptive evolution," Stoch. Models 24, 2008, pp. 2-44.

[9] Shashahani, S. "A new mathematical framework for the study of linkage and selection," Mem. Amer. Math. Soc., 17, 1979.

[10] Svirezhev, Y. M. "Optimum principles in genetics". In: Studies on Theoretical Genetics, V. A. Ratner (Ed), 1972, pp. 86-102, Novosibirsk, USSR Academy of Science.

[11] Hofbauer, J. "The selection mutation equation," J. Math. Biology, 1985, 23, pp. 41-53.

[12] Darroch, J.N. and Seneta, E. "On quasi-stationary distributions in absorbing discrete-time finite Markov chains," J. Appl. Prob., 1965, 2, 88–100.

[13] Martínez, S. "Quasi-stationary distributions for birth-death chains. Convergence radii and Yaglom limit," Cellular automata and cooperative systems (Les Houches, 1992), 1993, 491–505, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 396, Kluwer Acad. Publ., Dordrecht.

[14] Martínez, S. and Vares, M.E. "A Markov chain associated with the minimal quasi-stationary distribution of birth-death chains," J. Appl. Prob., 1995, 32, No. 1, pp. 25–38.

[1]Laboratoire de Physique Théorique et Modélisation, CNRS-UMR 8089 et Université de Cergy-Pontoise, 2 Avenue Adolphe Chauvin, 95302, Cergy-Pontoise, FRANCE, [2] Departamento de Ingeniería Matemática, Centro Modelamiento Matemático, UMI 2807, UCHILE-CNRS, Casilla 170-3 Correo 3, Santiago, CHILE, E-mail: Thierry.Huillet@u-cergy.fr and smartine@dim.uchile.cl