Information and (co-)variances in discrete evolutionary genetics involving solely selection

Thierry Huillet

INFORMATION AND (CO-)VARIANCES IN DISCRETE EVOLUTIONARY GENETICS INVOLVING SOLELY SELECTION

THIERRY E. HUILLET

Abstract. The purpose of this Note is twofold: First, we introduce the general formalism of evolutionary genetics dynamics involving fitnesses, under both the deterministic and stochastic setups, and chiefly in discrete-time. In the process, we particularize it to a one-parameter model where only a selection parameter is unknown. Then and in a parallel manner, we discuss the estimation problems of the selection parameter based on a single-generation frequency distribution shift under both deterministic and stochastic evolutionary dynamics. In the stochastics, we consider both the celebrated Wright-Fisher and Moran models.

Keywords: Evolutionary genetics, covariances, fitness landscape, selection.

Topics: Evolutionary processes (theory), Population dynamics (Theory).

1. Introduction and outline

In this Note, we revisit the basics of both the deterministic and stochastic dynamics arising in discrete-time evolutionary genetics (EG). We start with the haploid case with K alleles before switching to the more tricky diploid case. In the course of the exposition, we shall focus on a particular one-parameter selection instance of the general fitness model for which only the selection parameter is assumed to be unknown.

Let us summarize and comment the content of Section 2. In the deterministic haploid case, the updates of the allele frequency distributions are driven by the relative fitnesses of the alleles, ending up in a state where only the fittest will survive. From the dynamics, it appears that the mean fitness increases as time passes by, the rate of increase being the variance in relative fitness. This constitutes the core of the Fisher theorem of natural selection (FTNS). Introducing a discrete version of the Fisher information on time brought about by the allelic frequencies, it follows that one can identify this Fisher information with the variance in relative fitness.

In the deterministic diploid case, there is a similar updating dynamics but now on the full array of the genotype frequencies. When mating is random so that the Hardy-Weinberg law applies, we may look at the induced marginal allelic frequencies dynamics. It follows that the induced closed-form allelic updating dynamics looks quite similar to the one occurring in the haploid case except that the mean fitness now is a quadratic form in the current frequencies whereas marginal fitnesses no longer are constant but affine functions in these frequencies. In this context, the FTNS still holds true but, as a result of the fitness landscape being more complex, there is a possibility for a polymorphic equilibrium state to emerge. A short incursion in the continuous-time setting shows that here again one can identify the variance in relative fitness to a familiar Fisher information on time brought about by the frequencies. If one rather looks at the partial rate of increase of the mean fitness, one can identify it with half the allelic variance in relative fitness which is that component of the full genotypic variance in relative fitness which can be explained additively by the alleles constituting the genotypes, [3]. The remaining interaction part of this decomposition of the full genotypic variance can naturally be attributed to the dominance relationships between the alleles. In the interpretation of some authors, including W. J. Ewens ([3], p. $64-67$), this last property based on the partial rate of increase rather constitutes the essence of the FTNS. Using this circle of ideas, it follows that looking at the partial rate of increase of the mean fitness also makes sense when dealing with the full array of the genotype frequencies, no matter what form of mating is at stake. It also involves an allelic variance in relative fitness. In each case, we keep looking at the incarnation of these results when dealing with the one-parameter selection model. So far the results introduced and discussed can be found to be classical, our own contribution being perhaps to put things in order and fix the notations and formalism in a clear way. An excellent introduction to these and related problems can be found in [7]. We believe that the following developments can be considered as being entirely new.

To end up with Section 2, we discuss the estimation problem of the selection parameter based on both the current and shifted allele frequencies observations. Similarly, we discuss this problem when the observable is a general scalar output of the current frequency distribution. When looking at the updating of this output, we encounter a particular incarnation of the Price equation, [4].

Section 3 is devoted to the stochastic version of these considerations when the transitions in the constitutive allelic population sizes are given by a $K-$dimensional Wright-Fisher model with total constant-size (see [3] and [8]). We show that the mean of the increment of the random absolute mean fitness is positive, whereas its rate of increase differs from its variance. We suggest that when the size of the total allelic population goes to infinity, one should recover part of the marginal deterministic theory. In the selection example, we compute the classical Fisher information on the selection parameter and exhibit its possible use in the estimation problem. We finally present some comparative issues pertaining to a related model of fundamental importance in the context of stochastic EG: the Moran model.

Lots remain to be done in the same spirit, in particular including mutations and considering the multi-loci case with recombination. We again emphasize that in our models, there are no mutations included.

2. EG theory: the deterministic point of view

We start with the haploid case before moving to the diploid case, see ([3] and [7] for similar concerns).

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 on the $K-$simplex S_K is given by x_k, $k = 1, ...,K$. Let $\mathbf{x} := x_k$, $k = 1, ..., K$ stand for the column-vector[1] of these frequencies with $|\mathbf{x}| := \sum_k x_k = 1$. Let $w_k > 0$, $k = 1, ..., K$ be the absolute fitness of allele A_k. Let

$$\overline{w}_k(\mathbf{x}) = \frac{w_k}{w(\mathbf{x})} \tag{1}$$

be the relative fitness of allele A_k where $w(\mathbf{x}) := \sum_l w_l x_l$ represents the mean fitness of the population at time t. We shall also let

$$\sigma^2(\mathbf{x}) = \sum_{k=1}^{K} x_k (w_k - w(\mathbf{x}))^2 \tag{2}$$

stand for the variance in absolute fitness and

$$\overline{\sigma}^2(\mathbf{x}) = \sum_{k=1}^{K} x_k (\overline{w}_k(\mathbf{x}) - 1)^2 = \sigma^2(\mathbf{x}) / w(\mathbf{x})^2 \tag{3}$$

will be the variance in relative fitness.

From the deterministic EG point of view, the discrete-time update of the allele frequency distribution on the simplex S_K is given by[2]

$$x'_k = p_k(\mathbf{x}), \; k = 1, ..., K. \tag{4}$$

where $p_k(\mathbf{x}) := x_k \overline{w}_k(\mathbf{x})$. The quantity $\overline{w}_k(\mathbf{x}) - 1$ interprets as the frequency-dependent Malthus growth rate parameter of x_k.

The vector $\mathbf{p}(\mathbf{x}) := p_k(\mathbf{x})$, $k = 1, ..., K$, maps S_K into S_K. In vector form, with $\overline{\mathbf{w}}(\mathbf{x}) := \overline{w}_k(\mathbf{x})$, $k = 1, ..., K$ and $D_{\mathbf{x}}$:=diag$(x_k, k = 1, ..., K)$, the nonlinear deterministic EG dynamics reads:

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

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

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

Avoiding the trivial case where fitnesses are all equal, without loss of generality, we can assume that either $w_1 \geq ... \geq w_K = 1$ or $w_1 \leq ... \leq w_K = 1$. Thus allele A_1 or A_K has largest fitness. The deterministic EG dynamics attains an equilibrium where only the fittest will survive. The equilibrium is an extremal state of the boundary of S_K.

Example (selection). In general, all the w_k are unknown but sometimes, the set of unknowns can be reduced to 1 as follows: Let $s > -1$ stand for a selection parameter. Let a_k, $k = 1, ..., K$ stand for a known $[0, 1]-$valued decreasing sequence with $a_1 = 1$, $a_K = 0$ and assume $w_k = 1+sa_k$. The fitness landscape is $w(\mathbf{x}) = 1+sa(\mathbf{x})$ where $a(\mathbf{x}) := \sum_k a_k x_k$. A possible choice of a_k is $a_k = (K - k) / (K - 1)$ leading to equally spaced fitnesses with $w_{k+1} - w_k = -s/(K - 1)$. An alternative choice is

[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.

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

$a_k = (K/k-1)/(K-1)$ with $w_{k+1} - w_k = -\frac{Ks}{K-1}\frac{1}{k(k+1)}$. Depending on $s>0$ or $s<0$, the unit fitness 1 is either the minimal or the maximal value of the ordered w_ks. Although this particular model does not cover the class of all possible fitnesses, its generality is sufficient for our purposes and allows a considerable simplification of the exposition which otherwise would become tedious. It does not alter the general line of thought in a major way. ♦

According to the EG dynamical system (4), for each k, the relative fitness decreases as time passes by. Indeed, with $\Delta\overline{w}_k(\mathbf{x}) := \overline{w}_k(\mathbf{x}') - \overline{w}_k(\mathbf{x})$

$$\Delta\overline{w}_k(\mathbf{x}) = \overline{w}_k\left(D_{\overline{\mathbf{w}}(\mathbf{x})}\mathbf{x}\right) - \overline{w}_k(\mathbf{x}) = \frac{w_k}{w\left(D_{\overline{\mathbf{w}}(\mathbf{x})}\mathbf{x}\right)} - \frac{w_k}{w(\mathbf{x})}$$

$$= \frac{w_k}{\sum_l w_l\overline{w}_l(\mathbf{x})x_l} - \frac{w_k}{\sum_l w_l x_l} < 0$$

because $\overline{w}_l(\mathbf{x}) = \frac{w_l}{w(\mathbf{x})}$ and $w(\mathbf{x})^2 < \sum_l w_l^2 x_l$. However, unless the equilibrium state is attained, the absolute mean fitness $w(\mathbf{x})$ increases:

$$\Delta w(\mathbf{x}) = w(\mathbf{x}') - w(\mathbf{x}) = \sum_k w_k \Delta x_k$$

$$= \sum_k w_k x_k(\overline{w}_k(\mathbf{x}) - 1) = \frac{\sum_k w_k^2 x_k}{w(\mathbf{x})} - w(\mathbf{x}) > 0.$$

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

$$\frac{\Delta w(\mathbf{x})}{w(\mathbf{x})} = \sum_k x_k(\overline{w}_k(\mathbf{x}) - 1)^2 = \sum_k \frac{(\Delta x_k)^2}{x_k} \tag{5}$$

which is the variance in relative fitness $\overline{\sigma}^2(\mathbf{x})$ defined in (3). These last two facts are sometimes termed the 1930s Fisher fundamental theorem of natural selection (FTNS).

Remarks.

(i) The expression appearing in the right-hand side of (5) is also

$$\sum_k \frac{(\Delta x_k)^2}{x_k} = \sum_k x_k\left(\frac{\Delta x_k}{x_k}\right)^2.$$

The discrete frequency distribution $\mathbf{x}$ depends on the time parameter $t \in \{0,1,2,...\}$ which is itself discrete. The quantity $I_{\mathbf{x}}(t) := \sum_k x_k\left(\frac{\Delta x_k}{x_k}\right)^2$ may therefore be interpreted as a discrete version of the Fisher information about t brought by $\mathbf{x}$. From (5), we get that the rate of increase of the mean fitness (which is the variance in relative fitness) identifies with this Fisher information

$$\frac{\Delta w(\mathbf{x})}{w(\mathbf{x})} = \overline{\sigma}^2(\mathbf{x}) = I_{\mathbf{x}}(t) > 0. \text{ ♦} \tag{6}$$

(ii) When $w_k = 1 + sa_k$ as in the selection example, the variance in relative fitness reads

$$\overline{\sigma}^2(\mathbf{x}) = \left(\frac{s}{1+sa(\mathbf{x})}\right)^2 \sum_k x_k(a_k - a(\mathbf{x}))^2. \text{ ♦}$$

2.2. **Single locus: diploid population with K alleles.** We now run into similar considerations but with diploid populations whose genetical information governing their developments is carried by pairs of chromosomes. When considering the estimation problem, to avoid overburden notations that would blur the exposition, we shall limit ourselves to the special one-parameter fitness model where a single selection parameter s is unknown. Under this hypothesis, the estimation problem is over-simplified because it avoids estimating the full fitness array that would lead to additional notational and technical difficulties due to multidimensionality.

Joint EG dynamics. Let $w_{k,l} > 0$, $k, l = 1, ..., K$ stand for the absolute fitness of the genotypes $A_k A_l$ attached to a single locus. Assume $w_{k,l} = w_{l,k}$. Let W be the symmetric fitness matrix with $k, l-$entry $w_{k,l}$. Assume the current frequency distribution at time t of the genotypes $A_k A_l$ is given by $x_{k,l}$. Let X be the frequencies array with $k, l-$entry $x_{k,l}$. The joint EG dynamics in the diploid case is given by the updating:

$$x'_{k,l} = x_{k,l} \frac{w_{k,l}}{w(X)} \tag{7}$$

where the mean fitness w now is given by: $w(X) = \sum_{k,l} x_{k,l} w_{k,l}$. Define the relative fitness of the genotype $A_k A_l$ by: $\overline{w}_{k,l}(X) := \frac{w_{k,l}}{w(X)}$ and let $\overline{W}(X)$ be the matrix with entries $\overline{w}_{k,l}(X)$. Then the joint EG dynamics takes the matrix form:

$$X' = X \circ \overline{W}(X) = \overline{W}(X) \circ X$$

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

Let J be the $K \times K$ flat matrix whose entries are all 1. Then

$$\Delta X := X' - X = (X - J) \circ \overline{W}(X) = \overline{W}(X) \circ (X - J).$$

We shall also let

$$\sigma^2(X) = \sum_{k,l=1}^{K} x_{k,l} (w_{k,l} - w(X))^2 \tag{8}$$

stand for the genotypic variance in absolute fitness and

$$\overline{\sigma}^2(X) = \sum_{k,l=1}^{K} x_{k,l} (\overline{w}_{k,l}(X) - 1)^2 = \sigma^2(X) / w(X)^2 \tag{9}$$

will stand for the diploid variance in relative fitness.

Consider the problem of evaluating the increase of the mean fitness. We have

$$\Delta w(X) = \sum_{k,l} \Delta x_{k,l} w_{k,l} = \sum_{k,l} x_{k,l} \left(\frac{w_{k,l}^2}{w(X)} - w_{k,l} \right) = w(X) \overline{\sigma}^2(X) > 0 \tag{10}$$

with a relative rate of increase: $\Delta w(X) / w(X) = \overline{\sigma}^2(X)$. This is the full diploid version of the FTNS.

Marginal allelic dynamics. Assuming a Hardy-Weinberg equilibrium, the frequency distribution at time t, say $x_{k,l}$, of the genotypes $A_k A_l$ is given by: $x_{k,l} = x_k x_l$ where $x_k = \sum_l x_{k,l}$ is the marginal frequency of allele A_k in the whole genotypic population. The whole frequency information is now enclosed within $\mathbf{x} := x_k$, $k = 1, ..., K$. For instance, the mean fitness w now is given by the quadratic form:

$w(\mathbf{x}) = \sum_{k,l} x_k x_l w_{k,l} = \mathbf{x}^* W \mathbf{x}$ with $\mathbf{x}^*$ the transposed line vector of the column vector $\mathbf{x} = X\mathbf{1}$ ($\mathbf{1}$ the unit K-vector). We shall also let

$$\sigma^2(\mathbf{x}) = \sum_{k,l=1}^{K} x_k x_l (w_{k,l} - w(\mathbf{x}))^2 \tag{11}$$

stand for the genotypic variance in absolute fitness and

$$\overline{\sigma}^2(\mathbf{x}) = \sum_{k,l=1}^{K} x_k x_l (\overline{w}_{k,l}(\mathbf{x}) - 1)^2 = \sigma^2(\mathbf{x}) / w(\mathbf{x})^2$$

will stand for the diploid variance in relative fitness with $\overline{w}_{k,l}(\mathbf{x}) := w_{k,l}/w(\mathbf{x})$ the relative fitnesses. These quantities may now simply be indexed by $\mathbf{x}$.

Before we come to the diploid marginal EG dynamics, let us make the following remarks. Let

$$S^2(\boldsymbol{\alpha}) := \sum_{k,l=1}^{K} x_k x_l (w_{k,l} - w(\mathbf{x}) - \alpha_k - \alpha_l)^2 . \tag{12}$$

The values of $\boldsymbol{\alpha}$ minimizing $S^2(\boldsymbol{\alpha})$ are easily seen to be $\boldsymbol{\alpha}^* = \alpha_k^* = w_k(\mathbf{x}) - w(\mathbf{x})$, $k = 1, ..., K$. We shall let

$$\sigma_D^2(\mathbf{x}) := S^2(\boldsymbol{\alpha}^*) = \sum_{k,l=1}^{K} x_k x_l (w_{k,l} - (w_k(\mathbf{x}) + w_l(\mathbf{x}) - w(\mathbf{x})))^2 \tag{13}$$

and call it the dominance variance. Then we get

$$\sigma^2(\mathbf{x}) = \sigma_D^2(\mathbf{x}) + \sigma_A^2(\mathbf{x}) . \tag{14}$$

The variance $\sigma_A^2(\mathbf{x})$ is that component of the total variance in absolute fitness of the genotypes which can be explained additively by the alleles constituting those genotypes, [2]. Indeed, we can easily check that

$$\sigma_A^2(\mathbf{x}) = 2\sum_{k=1}^{K} x_k (w_k(\mathbf{x}) - w(\mathbf{x}))^2 . \tag{15}$$

The number $\sigma_A^2(\mathbf{x})/2$ can be interpreted in terms of the fitness covariance between parent and offspring in the updating step (see Ewens, [3], p. 7).

The residual part $\sigma_D^2(\mathbf{x})$ is that component of $\sigma^2(\mathbf{x})$ which can be explained by the interactions pertaining to dominance between the alleles forming the genotypes.

Consider now the update of the allelic marginal frequencies $\mathbf{x}$ themselves. 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 (4) by:

$$x_k' = x_k \overline{w}_k(\mathbf{x}) =: p_k(\mathbf{x}), \; k = 1, ..., K \tag{16}$$

where now: $\overline{w}_k(\mathbf{x}) := \frac{w_k(\mathbf{x})}{w(\mathbf{x})}$ is the relative fitness of A_k. In vector form

$$\mathbf{x}' = \frac{D_{\mathbf{x}} W \mathbf{x}}{\mathbf{x}^* W \mathbf{x}} = D_{\overline{\mathbf{w}}(\mathbf{x})} \mathbf{x},$$

where $\overline{\mathbf{w}}(\mathbf{x}) := \overline{w}_k(\mathbf{x})$, $k = 1, ..., K$. Again, the mean fitness $w(\mathbf{x})$, as a Lyapounov function, increases as time passes by. We indeed have

$$\Delta w(\mathbf{x}) = w(\mathbf{x}') - w(\mathbf{x}) = \sum_{k,l} x_k \overline{w}_k(\mathbf{x}) w_{k,l} x_l \overline{w}_l(\mathbf{x}) - \sum_{k,l} x_k w_{k,l} x_l > 0,$$

because, defining $0 < X(\mathbf{x}) := \sum_{k,l} x_k (1 - \overline{w}_k(\mathbf{x})) w_{k,l} (1 - \overline{w}_l(\mathbf{x})) x_l$, we have

$$\Delta w(\mathbf{x}) = X(\mathbf{x}) + \frac{2}{w(\mathbf{x})} \left(\sum_k x_k w_k(\mathbf{x})^2 - w(\mathbf{x})^2 \right) > 0.$$

Its partial rate of increase due to frequency shifts only is

$$\frac{\Delta_P w(\mathbf{x})}{w(\mathbf{x})} := \frac{\sum_k \Delta x_k w_k(\mathbf{x})}{w(\mathbf{x})}.$$

This quantity is half the allelic variance in relative fitness $\sigma_A^2(\mathbf{x}) / \left(2w(\mathbf{x})^2\right) = \overline{\sigma}_A^2(\mathbf{x})/2$. Indeed,

$$\frac{\Delta_P w(\mathbf{x})}{w(\mathbf{x})} = \sum_k x_k (\overline{w}_k(\mathbf{x}) - 1)^2 = \sum_k \frac{(\Delta x_k)^2}{x_k} = \overline{\sigma}_A^2(\mathbf{x})/2. \tag{17}$$

The mean fitness increase phenomena (either global or partial) occur till the EG dynamics reaches an equilibrium state. In the diploid case, this dynamics can have more complex equilibrium points, satisfying $w_k(\mathbf{x}_{eq}) = w_1(\mathbf{x}_{eq})$, $k = 2, ..., K$ and $\sum_l x_{eq,l} = 1$. In particular, a stable internal (polymorphic) equilibrium state can exist, a necessary and sufficient condition being that W has exactly one strictly positive dominant eigenvalue and at least one strictly negative eigenvalue (see Kingman, [6]) or else that the sequence of principal minors of W alternates in sign. An internal polymorphic equilibrium state is asymptotically stable iff it is an isolated local maximum of the mean fitness. If this is the case, there is a unique $\mathbf{z} > 0$ for which $W\mathbf{z} = \mathbf{1}$ and the equilibrium polymorphic state is $\mathbf{x}_{eq} = \mathbf{z}/|\mathbf{z}|$. Moreover, starting from any initial condition in the interior of S_K, all trajectories are attracted by this $\mathbf{x}_{eq}$. When there is no such unique globally stable polymorphic equilibrium, all trajectories will still converge but perhaps to a local equilibrium state where some alleles get extinct.

Except for the fact that the mean fitness now is a quadratic form in $\mathbf{x}$ and that the marginal fitness of A_k now is frequency-dependent, depending linearly on $\mathbf{x}$, as far as the marginal frequencies are concerned, the updating formalism (16) in the diploid case looks very similar to the one in (4) describing the haploid case. In the diploid case, assuming fitnesses are multiplicative, say with $W_{k,l} = w_k w_l$, then $\overline{w}_k(\mathbf{x}) := \frac{w_k(\mathbf{x})}{\mathbf{x}^* W \mathbf{x}} = \frac{w_k}{\sum_l w_l x_l}$ and the dynamics (16) boils down to (4). However, the mean fitness in this case is $w(\mathbf{x}) = \left(\sum_l w_l x_l\right)^2$ and not $\sum_l w_l x_l$ as in the haploid case.

Example (selection). In general the whole fitness matrix is unknown. In some cases, only one selection parameter s is to be determined (estimated from data). Assume indeed $w_{k,l} = 1 + s a_{k,l}$ where $a_{k,l} = a_{l,k} \in [0, 1]$ are known and $s > -1$. A natural choice could be $a_{k,l} = \frac{K-k}{K-1} \frac{K-l}{K-1}$, with $a_{1,1} = 1$ and $a_{K,K} = 0$. Or else:

$a_{k,l} = (K/k-1)(K/l-1)/(K-1)^2$. The simple popular model $a_{k,l} = \delta_{k,l}$ is also of wide use in this context (see [3], p. 53 or [7] p. 14).

Then $W = J + sA$ where A is a known matrix whose $k,l-$entry is $a_{k,l}$. With $a_k(\mathbf{x}) = \sum_l a_{k,l}x_l$ and $a(\mathbf{x}) = \mathbf{x}^* A\mathbf{x}$, the EG dynamics reads

$$x'_k = x_k \frac{1 + sa_k(\mathbf{x})}{1 + sa(\mathbf{x})},\ k = 1, ..., K. \tag{18}$$

The allelic variance in relative fitness reads

$$\overline{\sigma}_A^2(\mathbf{x})/2 = \left(\frac{s}{1 + sa(\mathbf{x})}\right)^2 \sum_k x_k (a_k(\mathbf{x}) - a(\mathbf{x}))^2. \ \blacklozenge \tag{19}$$

Remarks:

(i) There is an alternative vectorial representation of the dynamics (16) and (18). Define the symmetric positive-definite matrix $Q(\mathbf{x})$ with quadratic entries in the frequencies:

$$Q(\mathbf{x})_{k,l} = x_k(\delta_{k,l} - x_l).$$

Introduce the column vector of the relative fitnesses: $\overline{\mathbf{w}}(\mathbf{x}) = \overline{w}_k(\mathbf{x}),\ k = 1, ..., K$ (with $\overline{\mathbf{w}}(\mathbf{x}) =: \nabla V(\mathbf{x}) = \frac{1}{2}\nabla \log w(\mathbf{x})$, half the gradient of the logarithm of mean fitness). Then, (16) may be recast as the gradient-like replicator dynamics:

$$\Delta\mathbf{x} = Q(\mathbf{x})\overline{\mathbf{w}}(\mathbf{x}) = \frac{1}{w(\mathbf{x})}Q(\mathbf{x})W\mathbf{x} = Q(\mathbf{x})\nabla V(\mathbf{x}), \tag{20}$$

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

$$\nabla V(\mathbf{x})^* \Delta\mathbf{x} = \nabla V(\mathbf{x})^* Q(\mathbf{x})\nabla V(\mathbf{x}) \geq 0.$$

In the selection case when $w_k(\mathbf{x}) = 1 + sa_k(\mathbf{x})$, using $Q(\mathbf{x})\mathbf{1} = \mathbf{0}$:

$$\Delta\mathbf{x} = Q(\mathbf{x})(\overline{\mathbf{w}}(\mathbf{x}) - \mathbf{1}) = \frac{s}{1 + sa(\mathbf{x})}Q(\mathbf{x})A\mathbf{x}. \ \blacklozenge$$

(ii) Although we shall not run into details pertaining to the continuous-time setting, let us say a few words on this particular aspect. In continuous-time $t \geq 0$, the dynamics of $x_k := x_k(t)$ is

$$\overset{\cdot}{x}_k = x_k(w_k(\mathbf{x}) - w(\mathbf{x})),\ k = 1, ..., K$$

where the 'dot' is the time-derivative. The growth rate is driven by the average excess in mean fitness $w_k(\mathbf{x}) - w(\mathbf{x})$. Alternatively, the dynamics on S_K is $\overset{\cdot}{\mathbf{x}} = Q(\mathbf{x})W\mathbf{x}$ with $\frac{d}{dt}|\mathbf{x}| = 0$ because $|\mathbf{x}| = \mathbf{1}^*\mathbf{x} = \langle \mathbf{1}, \mathbf{x}\rangle$ and $\mathbf{1}^*Q(\mathbf{x}) = Q(\mathbf{x})\mathbf{1} = \mathbf{0}$.

In the special selection case, $\overset{\cdot}{\mathbf{x}} = sQ(\mathbf{x})A\mathbf{x}$ or

$$\overset{\cdot}{x}_k = sx_k(a_k(\mathbf{x}) - a(\mathbf{x})),\ k = 1, ..., K.$$

In this case, the positive quantity

$$\sum_{k=1}^K \frac{\left(\overset{\cdot}{x}_k\right)^2}{x_k} = \sum_{k=1}^K x_k \left[\frac{d(\log x_k)}{dt}\right]^2 = s^2 \sum_{k=1}^K x_k (a_k(\mathbf{x}) - a(\mathbf{x}))^2$$

may be viewed as the familiar Fisher information $I_{\mathbf{x}} := I_{\mathbf{x}}(t)$ of the frequency distribution $\mathbf{x}$, as a discrete probability distribution parameterized by continuous time t. One can check that, if the mean fitness is $w(\mathbf{x}) = 1 + sa(\mathbf{x})$, then

$$\overset{\cdot}{w}(\mathbf{x}) = s\overset{\cdot}{a}(\mathbf{x}) = 2I_{\mathbf{x}}.$$

So the time derivative of $w(\mathbf{x})$ coincides with twice this Fisher information. This constitutes the diploid continuous-time version of (6).

Defining the dimensionless parameter $\theta := st$ and looking at the time-changed frequencies $\pi_k(\theta) := x_k(\theta/s)$, we get the $s-$free dynamics

$$\overset{\cdot}{\pi}_k = \pi_k(a_k(\boldsymbol{\pi}) - a(\boldsymbol{\pi})), \ k = 1, ..., K,$$

where the 'dot' now is the derivative with respect to θ. Clearly, $\overset{\cdot}{w}(\boldsymbol{\pi}) = \overset{\cdot}{a}(\boldsymbol{\pi}) = 2I_{\boldsymbol{\pi}}(\theta) > 0$. ♦

Partial change of mean fitness. Let us return to the joint EG dynamics where no hypothesis on mating was made and consider the full mean fitness

$$w(X) = \sum_{k,l} x_{k,l} w_{k,l}. \tag{21}$$

Define $\alpha_k = w_k(X) - w(X)$ where $w_k(X) := \sum_l x_{k,l} w_{k,l}/x_k$ and $x_k := \sum_l x_{k,l}$ is the marginal frequency of A_k. Replace the expression (21) by the equally correct

$$w(X) = \sum_{k,l} x_{k,l}(w(X) + \alpha_k + \alpha_l), \tag{22}$$

suggesting that the fitnesses of the genotypes $A_k A_l$ would rather be $w_{k,l}^{(\alpha)} := w(X) + \alpha_k + \alpha_l$.

Define the partial change, say Δ_P, of mean fitness as

$$\Delta_P w(X) := \sum_{k,l} \Delta x_{k,l}(w(X) + \alpha_k + \alpha_l)$$

where only a variation in the frequency term is considered. After some elementary algebra, we get

$$\Delta_P w(X) = \sum_{k,l} \Delta x_{k,l}(\alpha_k + \alpha_l) = 2\sum_k \alpha_k \sum_l \Delta x_{k,l}$$

$$= 2\sum_k \alpha_k \Delta x_k =: \sigma_A^2(X)/w(X),$$

leading to a partial rate of increase $\Delta_P w(X)/w(X) = \overline{\sigma}_A^2(X)$ which is similar to (17). In the alternative Castilloux-Lessard interpretation of this phenomenon, [1], defining $\Delta x_{k,l}^{(\alpha)} := \Delta x_{k,l} \frac{\alpha_k + \alpha_l}{w(X)}$ and observing

$$\Delta_P w(X) = \sum_{k,l} \Delta x_{k,l}^{(\alpha)} w_{k,l},$$

the partial change involves an allele-based modification of the genotype frequencies while the genotype fitnesses $w_{k,l}$ are kept unchanged.

2.3. **Estimation of s.** We now switch to the announced estimation of s problem which seems to be new.

Assume the current and updated frequencies $\mathbf{x}$ and $\mathbf{x}'$ are being observed at some times t, $t+1$. We wish to use this information to estimate the unknown value of s. Let first s_k^* be the estimate of s which explains the observable $(\mathbf{x};x_k')$ at best. Clearly, from (18),

$$s_k^* = \frac{x_k' - x_k}{a_k(\mathbf{x})\,x_k - x_k' a(\mathbf{x})} \tag{23}$$

does the job. A natural estimator $s^* = s^*(\mathbf{x};\mathbf{x}')$ of s which explains best the observable $(\mathbf{x};\mathbf{x}')$ is:

$$s^* = \arg\min_s \sum_k x_k x_k' \left(s - s_k^*\right)^2$$

which is:

$$s^* = \frac{1}{\sum_l x_l x_l'} \sum_k x_k x_k' s_k^*, \tag{24}$$

a weighted average of the s_k^* attributing more credit to s_k^* when $x_k x_k'$ is largest.

Sometimes, the x_k, x_k' are not directly observed. Rather, what is observed is the scalar output:

$$y := h(\mathbf{x}) := \sum_k x_k h_k(\mathbf{x})$$

for some known family of measurements $h_k(\mathbf{x})$, $k = 1,...,K$ given the process $\mathbf{x}$ is in state k. Simple but important examples are $h_k(\mathbf{x}) = x_k^{\alpha-1}$ $(\alpha > 1)$ in which case, $y = h(\mathbf{x}) = \sum_k x_k^\alpha$ is $\alpha-$homozygosity (typically $\alpha = 2$), or $h_k(\mathbf{x}) = -\log x_k$ in which case, $y = h(\mathbf{x}) = -\sum_k x_k \log x_k$ is the Shannon entropy of the frequency distribution.

Let κ be a discrete random variable with $\mathbf{P}(\kappa = k) = x_k$, $k = 1,...,K$ so that $y := \mathbf{E}(h_\kappa(\mathbf{x}))$, the mathematical expectation with respect to κs law. From (18), we have:

$$\begin{aligned} y' = h(\mathbf{x}') &:= \sum_k x_k' h_k(\mathbf{x}') = \sum_k \Delta x_k h_k(\mathbf{x}') + \sum_k x_k h_k(\mathbf{x}') \\ &= \frac{s}{1+sa(\mathbf{x})} \sum_k x_k h_k(\mathbf{x}')\left(a_k(\mathbf{x}) - a(\mathbf{x})\right) + \mathbf{E}\left(h_\kappa(\mathbf{x}')\right) \\ &= \frac{s}{1+sa(\mathbf{x})} \mathbf{Cov}\left(h_\kappa(\mathbf{x}'), a_\kappa(\mathbf{x})\right) + \mathbf{E}\left(h_\kappa(\mathbf{x}')\right). \end{aligned}$$

Therefore, the observed shift in the measurement is

$$\Delta y = \frac{s}{1+sa(\mathbf{x})} \mathbf{Cov}\left(h_\kappa(\mathbf{x}'), a_\kappa(\mathbf{x})\right) + \mathbf{E}\left(\Delta h_\kappa(\mathbf{x})\right) \tag{25}$$

and an estimate s^* based on Δy can immediately be written down by mere solving the above equation (25) which is reminiscent of a Price equation (see [4]). It involves two terms, one which is related to the correlation between the measurement at $t+1$ and the fitness function at t that has to do with the frequency shift only, another related to the induced change of the character value only.

3. EG theory: the stochastic point of view

With $\overline{w}_k(\mathbf{x}) := \frac{w_k(\mathbf{x})}{w(\mathbf{x})}$, let $p_k(\mathbf{x}) := x_k\overline{w}_k(\mathbf{x})$, $k = 1, ..., K$ with $\sum_k p_k(\mathbf{x}) = 1$ be defined as in the previous Section either from allelic or genotypic fitnesses. We shall assume throughout that the special selection model assumptions: $w_k(\mathbf{x}) = 1 + sa_k(\mathbf{x})$ and $w(\mathbf{x}) = 1 + sa(\mathbf{x})$ are at stake.

3.1. **The Wright-Fisher model.** We start considering similar problems under the Wright-Fisher model.

The Model and first properties. Consider an allelic population with constant size N. In the haploid (diploid) case, N is (twice) the number of real individuals. Let $\mathbf{i} := i_k$ and $\mathbf{i}' := i'_k$, $k = 1, ..., K$ be two vectors of integers quantifying the size of the allelic populations at two consecutive generations t and $t+1$. With $|\mathbf{i}| = \sum_k i_k$, therefore $\left|\frac{\mathbf{i}}{N}\right| = \left|\frac{\mathbf{i}'}{N}\right| = 1$ on S_K . Suppose the stochastic EG dynamics now is given by a Markov chain whose one-step transition matrix P from states $\mathbf{I} = \mathbf{i}$ to $\mathbf{I}' = \mathbf{i}'$ is given by the multinomial Wright-Fisher (WF) model

$$\mathbb{P}\left(\mathbf{I}'_{t+1} = \mathbf{i}' \mid \mathbf{I}_t = \mathbf{i}\right) =: P(\mathbf{i}, \mathbf{i}') = \binom{N}{i'_1 \cdots i'_K} \prod_{k=1}^{K} p_k\left(\frac{\mathbf{i}}{N}\right)^{i'_k}. \tag{26}$$

The state-space dimension of this Markov chain is $\binom{N+1}{K-1}$ (the number of compositions of integer N into K non-negative parts).

Let $\mathbf{e}_l$ be the $K-$null vector except for its $l-$th entry which is 1. The extremal states $S_K^* := \{\mathbf{i}_l^* := N\mathbf{e}_l, l = 1, ..., K\}$, are all absorbing for this Markov chain because $p_k\left(\frac{\mathbf{i}_l^*}{N}\right) = \delta_{k,l}$. Under our assumptions, the chain is not recurrent. Depending on the initial condition, say $\mathbf{i}_0$, the chain will necessarily end up in one of the extremal states $\mathbf{i}_l^*$, with some probability, say $\pi_l(\mathbf{i}_0)$, which can be computed as follows. Let $\boldsymbol{\pi}_l := \pi_l(\mathbf{i})$, $\mathbf{i} \in S_K$ be an harmonic function of the WF Markov chain, solution to:

$$(P - I)\boldsymbol{\pi}_l = \mathbf{0} \text{ if } \mathbf{i} \in S_K \backslash S_K^* \text{ and } \boldsymbol{\pi}_l = 1(\mathbf{i} = N\mathbf{e}_l) \text{ if } \mathbf{i} \in S_K^*. \tag{27}$$

It satisfies

$$\mathbb{P}(\mathbf{I}_\tau = N\mathbf{e}_l \mid \mathbf{I}_0 = \mathbf{i}_0) = \pi_l(\mathbf{i}_0),$$

where τ ($< \infty$ almost surely) is the random hitting time of S_K^* for $\mathbf{I}_t$ and the $\pi_l(\mathbf{i}_0)$s are normalized so as $\sum_l \pi_l(\mathbf{i}_0) = 1$. Thus $\pi_l(\mathbf{i}_0)$ is the searched probability to end up in state $N\mathbf{e}_l$ starting from state $\mathbf{i}_0$. In the same vein, the expected hitting time $\alpha(\mathbf{i}_0) := \mathbb{E}_{\mathbf{i}_0}(\tau)$ solves:

$$\begin{aligned} (P - I)\boldsymbol{\alpha} &= \mathbf{1}, \mathbf{i}/N \in S_K \backslash S_K^* \\ \boldsymbol{\alpha} &= 0, \mathbf{i}/N \in S_K^* \end{aligned}$$

where $\boldsymbol{\alpha} := \alpha(\mathbf{i})$, $\mathbf{i} \in S_K$. With $\boldsymbol{\pi}_l$ the solution to the above Dirichlet problem, the equilibrium measure of the chain therefore is:

$$\pi_{eq} := \sum_{l=1}^{K} \pi_l(\mathbf{i}_0)\, \delta_{\mathbf{i}_l^*},$$

which depends on $\mathbf{i}_0$. Unless some prior information on $\mathbf{i}_0$ is given, we may assume that $\mathbf{i}_0 = \frac{N}{K}\mathbf{1}$ in which case one expects $\pi_l(\mathbf{i}_0) = 1/K$ and π_{eq} is uniform on the extremal states.

Necessarily, one allele will fixate and there is no polymorphic equilibrium state even when dealing with diploid populations. Which allele and with what probability will depend on the initial condition. Thanks to fluctuations, the picture therefore looks very different from the one pertaining to the deterministic theory. For analogies of this construction with statistical physics, see [10], [9].

The marginal transition matrix from $\mathbf{i}$ to $I'_k = i'_k$ is binomial $\text{bin}\big(N, p_k\left(\frac{\mathbf{i}}{N}\right)\big)$:

$$P\left(\mathbf{i}, i'_k\right) = \binom{N}{i'_k} p_k\left(\frac{\mathbf{i}}{N}\right)\left(1 - p_k\left(\frac{\mathbf{i}}{N}\right)\right)^{N-i'_k}.$$

With $p_k\left(\frac{\mathbf{i}}{N}\right) := \frac{i_k}{N}\overline{w}_k\left(\frac{\mathbf{i}}{N}\right)$, given $\mathbf{I} = \mathbf{i}$, the $k-$th component I'_k of the updated state is now random with:

$$\mathbb{E}_{\mathbf{i}}\left(I'_k\right) = Np_k\left(\frac{\mathbf{i}}{N}\right) \text{ and } \sigma^2_{\mathbf{i}}\left(I'_k\right) = Np_k\left(\frac{\mathbf{i}}{N}\right)\left(1 - p_k\left(\frac{\mathbf{i}}{N}\right)\right).$$

Mean fitness. We shall introduce the random increment in absolute mean fitness as

$$\Delta w_{\mathbf{I}'}\left(\frac{\mathbf{i}}{N}\right) := \sum_{k=1}^{K}\left(\frac{I'_k}{N} - \frac{i_k}{N}\right) w_k\left(\frac{\mathbf{i}}{N}\right) = s\sum_{k=1}^{K}\left(\frac{I'_k}{N} - \frac{i_k}{N}\right) a_k\left(\frac{\mathbf{i}}{N}\right). \tag{28}$$

Dropping for notational ease the argument $\frac{\mathbf{i}}{N}$ appearing in $\Delta w_{\mathbf{I}'}$, a_k, a and p_k, we get

$$\mathbb{E}_{\mathbf{i}}\Delta w_{\mathbf{I}'} = s\sum_{k=1}^{K}\frac{i_k}{N}\left(\frac{1+sa_k}{1+sa} - 1\right)a_k = \frac{s^2}{1+sa}\sum_{k=1}^{K}\frac{i_k}{N}\left(a_k - a\right)a_k$$

$$= \frac{s^2}{1+sa}\left[\sum_{k=1}^{K}\frac{i_k}{N}a_k^2 - a^2\right] > 0. \tag{29}$$

The mean of the increment of the random absolute mean fitness is positive (a random version of the FTNS). Its rate of increase is

$$\frac{\mathbb{E}_{\mathbf{i}}\Delta w_{\mathbf{I}'}\left(\frac{\mathbf{i}}{N}\right)}{w\left(\frac{\mathbf{i}}{N}\right)} = \left(\frac{s}{1+sa}\right)^2\left[\sum_{k=1}^{K}\frac{i_k}{N}a_k^2 - a^2\right], \tag{30}$$

involving the variance of the a_ks under the current frequency distribution $\frac{i_k}{N}$, $k = 1, ..., K$.

Let us now compute the variance of $\Delta w_{\mathbf{I}'}\left(\frac{\mathbf{i}}{N}\right)$. We get:

$$\sigma^2_{\mathbf{i}}\left(\Delta w_{\mathbf{I}'}\right) = s^2\sigma^2_{\mathbf{i}}\left(\sum_{k=1}^{K}\frac{I'_k}{N}a_k - a\right) = s^2\sigma^2_{\mathbf{i}}\left(\sum_{k=1}^{K}\frac{I'_k}{N}a_k\right) \tag{31}$$

$$= s^2\left[\mathbb{E}_{\mathbf{i}}\left(\sum_{k,k'=1}^{K}\frac{I'_kI'_{k'}}{N^2}a_ka_{k'}\right) - \left(\mathbb{E}_{\mathbf{i}}\left(\sum_{k=1}^{K}\frac{I'_k}{N}a_k\right)\right)^2\right].$$

It is proportional to the variance of the weighted outcomes $\sum_{k=1}^{K}\frac{I_k'}{N}a_k$ given $\mathbf{i}$. Using $\mathbb{E}_{\mathbf{i}}\left(I_k'I_{k'}'\right)=N\left(N-1\right)p_kp_{k'}$ and $\mathbb{E}_{\mathbf{i}}\left(I_k'^2\right)=Np_k+N\left(N-1\right)p_k^2$, we get

$$\sigma_{\mathbf{i}}^2\left(\Delta w_{\mathbf{I}'}\right)=s^2\left[\sum_{k,k'=1}^{K}\frac{N\left(N-1\right)p_kp_{k'}}{N^2}a_ka_{k'}+N\sum_{k=1}^{K}\frac{p_k}{N^2}a_k^2-\left(\sum_{k=1}^{K}p_ka_k\right)^2\right]$$

and so,

$$\sigma_{\mathbf{i}}^2\left(\Delta w_{\mathbf{I}'}\left(\frac{\mathbf{i}}{N}\right)\right)=\frac{s^2}{N}\left[\sum_{k=1}^{K}p_ka_k^2-\left(\sum_{k=1}^{K}p_ka_k\right)^2\right], \tag{32}$$

again involving the variance of the a_ks but now under the updated mean frequency distribution $p_k=\mathbb{E}_i\left(\frac{I_k'}{N}\right)=\frac{i_k}{N}\frac{1+sa_k}{1+sa}$, $k=1,...,K$.

We conclude that

$$\frac{\mathbb{E}_{\mathbf{i}}\Delta w_{\mathbf{I}'}\left(\frac{\mathbf{i}}{N}\right)}{w\left(\frac{\mathbf{i}}{N}\right)}\propto\sigma_{\mathbf{i}}^2\left(\Delta w_{\mathbf{I}'}\left(\frac{\mathbf{i}}{N}\right)\right) \tag{33}$$

as one might have expected from the analogies with the deterministic theory.

In fact, the full law of $\Delta w_{\mathbf{I}'}\left(\frac{\mathbf{i}}{N}\right)$ can be computed and the large N population limit is worth investigating. Indeed, its Laplace-Stieltjes transform (LST) reads

$$\mathbb{E}_{\mathbf{i}}\left(e^{-\lambda\Delta w_{\mathbf{I}'}\left(\frac{\mathbf{i}}{N}\right)}\right)=e^{\lambda sa}\mathbb{E}_{\mathbf{i}}\left(e^{-\frac{\lambda s}{N}\sum_{k=1}^{K}I_k'a_k}\right)=\left(\sum_{k=1}^{K}p_ke^{-\frac{\lambda s}{N}\left(a_k-a\right)}\right)^N,$$

suggesting from large deviation theory that, if $i_k:=\lfloor Nx_k\rfloor$, $k=1,...,K$

$$\Delta w_{\mathbf{I}'}\left(\frac{\lfloor N\mathbf{x}\rfloor}{N}\right)\underset{N\uparrow\infty}{\overset{a.s.}{\to}}\Delta w_{\mathbf{I}'}\left(\mathbf{x}\right)=s\sum_{k=1}^{K}p_k\left(\mathbf{x}\right)\left(a_k\left(\mathbf{x}\right)-a\left(\mathbf{x}\right)\right)$$

$$=\frac{s^2}{1+sa\left(\mathbf{x}\right)}\left(\sum_{k=1}^{K}x_ka_k\left(\mathbf{x}\right)^2-a\left(\mathbf{x}\right)^2\right),$$

which is the deterministic value $\sigma_A^2\left(\mathbf{x}\right)/\left(2\left(1+sa\left(\mathbf{x}\right)\right)\right)$ of the marginal deterministic theory (15).

Statistics. We now suppose the WF Markov chain is in state $\mathbf{i}$, with $\mathbf{i}\neq\mathbf{i}_l^*$ so that it has not yet reached any of its equilibrium states. Based on the observation $\mathbf{i}$, we would like to design estimators of the selection parameter s.

The log-likelihood of the model (26) is

$$\log P\left(\mathbf{i},\mathbf{i}'\right)=\log\binom{N}{i_1'...i_K'}+\sum_{k=1}^{K}i_k'\log\left[\frac{i_k}{N}w_k\left(\frac{\mathbf{i}}{N}\right)/w\left(\frac{\mathbf{i}}{N}\right)\right].$$

If $w_k=1+sa_k$ and $w=1+sa$, its derivative with respect to s is

$$\partial_s\log P\left(\mathbf{i},\mathbf{i}'\right)=\partial_s\left(\sum_{k=1}^{K}i_k'\log\left(1+sa_k\right)-N\log\left(1+sa\right)\right)$$

$$= \sum_{k=1}^{K} i'_k \frac{a_k}{1+sa_k} - N \frac{a}{1+sa}.$$

The value $s^{MLE} = s^{MLE}\left(\frac{\mathbf{i}}{N}, \frac{\mathbf{i}'}{N}\right)$ for which $\partial_s \log P(\mathbf{i}, \mathbf{i}') = 0$ is the Maximum Likelihood Estimator of s given the observable $(\mathbf{i}; \mathbf{i}')$. It is given by the implicit equation

$$\sum_{k=1}^{K} \frac{i'_k}{N} \frac{a_k\left(\frac{\mathbf{i}}{N}\right)}{1+s^{MLE} a_k\left(\frac{\mathbf{i}}{N}\right)} = \frac{a\left(\frac{\mathbf{i}}{N}\right)}{1+s^{MLE} a\left(\frac{\mathbf{i}}{N}\right)}. \tag{34}$$

It is probably biased. Let us compute the Fisher information on s enclosed in the observation $(\mathbf{i}; \mathbf{I}')$ which is

$$I_{\mathbf{i}}(s) = \mathbb{E}_{\mathbf{i}}\left[\left(\partial_s \log P(\mathbf{i}, \mathbf{I}')\right)^2\right]. \tag{35}$$

We get

$$I_{\mathbf{i}}(s) = \sum_{\mathbf{i}'} P(\mathbf{i}, \mathbf{i}') \left[\sum_{k=1}^{K} i'_k \frac{a_k}{1+sa_k} - \frac{Na}{1+sa}\right]^2 = \sigma_{\mathbf{i}}^2 \left(\sum_{k=1}^{K} I'_k \frac{a_k}{1+sa_k}\right)$$

$$= \sum_{k,k'=1}^{K} \frac{a_k}{1+sa_k} \frac{a_{k'}}{1+sa_{k'}} \mathbb{E}_{\mathbf{i}}\left(I'_k I'_{k'}\right) - \left(\frac{Na}{1+sa}\right)^2$$

and therefore

$$I_{\mathbf{i}}(s) = \frac{N}{1+sa}\left[\sum_{k=1}^{K} \frac{i_k}{N}\left(\frac{a_k^2}{1+sa_k} - \frac{a^2}{1+sa}\right)\right]. \tag{36}$$

The Fisher information is exactly the variance of the weighted outcomes $\sum_{k=1}^{K} \frac{I'_k}{N} \frac{a_k}{1+sa_k}$ given $\mathbf{i}$. We conclude that

$$\frac{\mathbb{E}_{\mathbf{i}} \Delta w_{\mathbf{I}'}\left(\frac{\mathbf{i}}{N}\right)}{w} \nsim I_{\mathbf{i}}(s). \tag{37}$$

From the expression of the mean $\mathbb{E}_{\mathbf{i}}(I'_k)$

$$\mathbb{E}_{\mathbf{i}}\left(\frac{I'_k}{N}\right) = \frac{i_k}{N} \frac{1+sa_k}{1+sa}$$

and, with $\langle i'_k \rangle_{\mathbf{i}} = \frac{1}{n}\sum_{m=1}^{n} i'_{k,m}$ the empirical average given $\mathbf{i}$ of I'_k based on a size-n sample observation of I'_k

$$s_k^* = \frac{\langle i'_k \rangle_{\mathbf{i}} - i_k}{a_k\left(\frac{\mathbf{i}}{N}\right) i_k - \langle i'_k \rangle_{\mathbf{i}} a\left(\frac{\mathbf{i}}{N}\right)}$$

is a first moment estimator of s explaining best $\left(\mathbf{i}; i'_{k,m}, m = 1, ..., n\right)$ and

$$s^* = \frac{1}{\sum_l i_l \langle i'_l \rangle_{\mathbf{i}}} \sum_k i_k \langle i'_k \rangle_{\mathbf{i}} s_k^*, \tag{38}$$

a moment estimator of s explaining best $\left(\mathbf{i}; \mathbf{i}'_m, m = 1, ..., n\right)$ whenever we are in possession of n observed copies $\mathbf{i}'$ of $\mathbf{I}'$ based on the same $\mathbf{i}$. This estimator is also biased.

With $\widehat{s}=s^{*}$ or s^{MLE}, let

$$R_{\mathbf{i}}\left(\widehat{s},s\right)=\mathbb{E}_{\mathbf{i}}\left[\left(\widehat{s}-s\right)^{2}\right]=\sigma_{\mathbf{i}}^{2}\left(\widehat{s}\right)+\left(\mathbb{E}_{\mathbf{i}}\left(\widehat{s}\right)-s\right)^{2} \tag{39}$$

be the quadratic risk function associated with the estimator $\widehat{s}$. By the Fréchet-Darmois-Cramer Rao theorem, we have

$$R_{\mathbf{i}}\left(\widehat{s},s\right)\geq\frac{1}{I_{\mathbf{i}}\left(s\right)}, \tag{40}$$

where, in this classical interpretation of the Fisher information, $I_{\mathbf{i}}\left(s\right)^{-1}$ appears as a universal lower bound of the estimator quadratic error.

Finally, we would like to stress that these considerations are also relevant in the context of another fundamental stochastic model arising in the context of evolutionary genetics. We shall give some elements of how to proceed with this model presenting very different properties.

3.2. **The $K-$allele Moran model.** We now focus on the estimation problem under the Moran model.

The Model. Let $\alpha,\beta\in\{1,...,K\}$. In the Moran version of the stochastic evolution, given $\mathbf{I}_{t}=\mathbf{I}=\mathbf{i}$, the only accessible values of $\mathbf{I}'$ are the neighboring states: $\mathbf{i}_{\alpha,\beta}':=\mathbf{i}+\mathbf{d}_{\alpha,\beta}$ where $\mathbf{d}_{\alpha,\beta}=(0,..,0,-1,0,...,1,0,...,0)$. Here -1 is in position α and 1 in position $\beta\neq\alpha$ corresponding to the transfer of an individual from cell α to cell β. The Moran stochastic EG dynamics now is given by a Markov chain whose one-step transition matrix P from states $\mathbf{I}=\mathbf{i}$ to $\mathbf{I}'=\mathbf{i}'$ is:

$$\mathbb{P}\left(\mathbf{I}_{t+1}=\mathbf{i}'\mid\mathbf{I}_{t}=\mathbf{i}\right)=0\text{ if }\mathbf{i}'\neq\mathbf{i}_{\alpha,\beta}'\text{ and} \tag{41}$$

$$\mathbb{P}\left(\mathbf{I}_{t+1}=\mathbf{i}_{\alpha,\beta}'\mid\mathbf{I}_{t}=\mathbf{i}\right)=:P\left(\mathbf{i},\mathbf{i}_{\alpha,\beta}'\right)=\frac{i_{\alpha}}{N}p_{\beta}\left(\frac{\mathbf{i}}{N}\right),$$

where $p_{\beta}\left(\frac{\mathbf{i}}{N}\right)$ is given by $p_{\beta}\left(\frac{\mathbf{i}}{N}\right):=\frac{i_{\beta}}{N}\overline{w}_{\beta}\left(\frac{\mathbf{i}}{N}\right)$.

Summing $P\left(\mathbf{i},\mathbf{i}_{\alpha,\beta}'\right)$ over α,β, $\beta\neq\alpha$ in (41), we get the holding probability

$$\mathbb{P}\left(\mathbf{I}_{t+1}=\mathbf{i}\mid\mathbf{I}_{t}=\mathbf{i}\right)=1-\sum_{\alpha,\beta:\beta\neq\alpha}\frac{i_{\alpha}}{N}p_{\beta}\left(\frac{\mathbf{i}}{N}\right)=\sum_{\alpha}\frac{i_{\alpha}}{N}p_{\alpha}\left(\frac{\mathbf{i}}{N}\right),$$

completing the characterization of the $K-$allele Moran model. The probability that in a one-step transition, the size of allele A_{α} population shrinks of one unit is:

$$\sum_{\beta\neq\alpha}P\left(\mathbf{i},\mathbf{i}_{\alpha,\beta}'\right)=\frac{i_{\alpha}}{N}\left(1-p_{\alpha}\left(\frac{\mathbf{i}}{N}\right)\right).$$

The probability that in a one-step transition, the size of allele A_{β} population undergoes a one unit growth is:

$$\sum_{\alpha\neq\beta}P\left(\mathbf{i},\mathbf{i}_{\alpha,\beta}'\right)=\left(1-\frac{i_{\beta}}{N}\right)p_{\beta}\left(\frac{\mathbf{i}}{N}\right).$$

As a nearest-neighbor random walk model, the Moran model has a much simpler transition matrix P of the Jacobi type. The equilibrium measure of the chain again is:

$$\pi_{eq} := \sum_{l=1}^{K} \pi_l \left(\mathbf{i}_0\right) \delta_{\mathbf{i}_l^*}, \tag{42}$$

where $\boldsymbol{\pi}_l$ again solves the Dirichlet problem (27) but with this new simpler Jacobi P.

In what follows, we assume the special one-parameter selection model leading to:

$$p_\beta \left(\frac{\mathbf{i}}{N}\right) = \frac{i_\beta}{N} \frac{1 + sa_\beta}{1 + sa}.$$

Mean fitness. Let us compute the LST of $\sum_k a_k I_k'$ in the context of a Moran model. We get the factorized form:

$$\begin{aligned}
\mathbb{E}_{\mathbf{i}} \left(e^{-\lambda \sum_k a_k I_k'}\right) &= \sum_{\alpha,\beta:\alpha\neq\beta} e^{-\lambda \sum_k a_k i'_{\alpha,\beta}(k)} P\left(\mathbf{i}, \mathbf{i}'_{\alpha,\beta}\right) + e^{-\lambda \sum_k a_k i_k} \sum_\beta \frac{i_\beta}{N} p_\beta \\
&= e^{-\lambda \sum_k a_k i_k} \left(\sum_{\alpha,\beta:\alpha\neq\beta} e^{-\lambda \sum_k a_k \mathbf{d}_{\alpha,\beta}(k)} P\left(\mathbf{i}, \mathbf{i}'_{\alpha,\beta}\right) + \sum_\beta \frac{i_\beta}{N} p_\beta\right) \\
&= e^{-\lambda \sum_k a_k i_k} \left(\sum_{\alpha,\beta:\alpha\neq\beta} e^{-\lambda(a_\beta - a_\alpha)} \frac{i_\alpha}{N} p_\beta + \sum_\beta \frac{i_\beta}{N} p_\beta\right) \\
&= e^{-\lambda \sum_k a_k i_k} \left(\sum_\beta e^{-\lambda a_\beta} p_\beta \sum_{\alpha\neq\beta} \frac{i_\alpha}{N} e^{\lambda a_\alpha} + \sum_\beta \frac{i_\beta}{N} p_\beta\right) \\
&= e^{-\lambda \sum_k a_k i_k} \left(\sum_\beta e^{-\lambda a_\beta} p_\beta \left(\sum_\alpha \frac{i_\alpha}{N} e^{\lambda a_\alpha} - \frac{i_\beta}{N} e^{\lambda a_\beta}\right) + \sum_\beta \frac{i_\beta}{N} p_\beta\right) \\
&= \left(e^{-\lambda \sum_k a_k i_k}\right) \left(\sum_\alpha \frac{i_\alpha}{N} e^{\lambda a_\alpha}\right) \left(\sum_\beta e^{-\lambda a_\beta} p_\beta\right).
\end{aligned}$$

Recalling $\Delta w_{\mathbf{I}'} \left(\frac{\mathbf{i}}{N}\right) = s \sum_{k=1}^{K} \left(\frac{I_k'}{N} - \frac{i_k}{N}\right) a_k$, this leads in particular to (compare with (29)):

$$\begin{aligned}
\mathbb{E}_{\mathbf{i}} \left(\Delta w_{\mathbf{I}'}\right) &= \frac{s}{N} \left(\mathbb{E}_{\mathbf{i}} \left(\sum_{k=1}^{K} I_k' a_k\right) - \sum_{k=1}^{K} i_k a_k\right) = \frac{s}{N} \left(\sum_\beta a_\beta \left(p_\beta - \frac{i_\beta}{N}\right)\right) \\
&= \frac{s^2}{N(1 + sa)} \left(\sum_\beta \frac{i_\beta}{N} a_\beta^2 - a^2\right) > 0.
\end{aligned} \tag{43}$$

The variance of $\Delta w_{\mathbf{I}'}\left(\frac{\mathbf{i}}{N}\right)$ could easily be computed. Using the above result, we indeed get:

$$\mathbb{E}_{\mathbf{i}}\left(e^{-\sum_l \lambda_l I_l'}\right)=e^{-\sum_l \lambda_l i_l}\sum_{\alpha}\frac{i_\alpha}{N}e^{\lambda_\alpha}\sum_{\beta}e^{-\lambda_\beta}p_\beta,$$

giving the joint LST of $\mathbf{I}'$ given $\mathbf{I}=\mathbf{i}$. Putting $\lambda_l=0$ if $l\neq k$, the $k^{th}-$marginal reads:

$$\mathbb{E}_{\mathbf{i}}\left(e^{-\lambda_k I_k'}\right)=e^{-\lambda_k i_k}\left(1-\frac{i_k}{N}+e^{\lambda_k}\frac{i_k}{N}\right)\left(1-p_k+e^{-\lambda_k}p_k\right)$$

which is of the random walk type. Indeed, we get: $\mathbb{P}_{\mathbf{i}}\left(I_k'=i_k'\right)=0$ if $i_k'\neq i_k\pm 1$ or $i_k'\neq i_k$ and

$$\mathbb{P}_{\mathbf{i}}\left(I_k'=i_k\right)=\left(1-\frac{i_k}{N}\right)\left(1-p_k\right)+\frac{i_k}{N}p_k$$

$$\mathbb{P}_{\mathbf{i}}\left(I_k'=i_k+1\right)=\left(1-\frac{i_k}{N}\right)p_k\ ;\ \mathbb{P}_{\mathbf{i}}\left(I_k'=i_k-1\right)=\frac{i_k}{N}\left(1-p_k\right).$$

In the special one-parameter case, we have

$$\mathbb{E}_{\mathbf{i}}\left(I_k'\right)=i_k+\left(p_k-\frac{i_k}{N}\right)=i_k+\frac{i_k}{N}\left(\frac{1+sa_k}{1+sa}-1\right).$$

Using previously introduced notations, this gives a first moment estimator of s explaining best $\left(\mathbf{i};i_{k,m}', m=1,...,n\right)$ as:

$$s_k^*=\frac{N\left(\langle i_k'\rangle_{\mathbf{i}}-i_k\right)}{i_k\left(a_k-a\right)}. \tag{44}$$

4. Concluding Remarks

In this Note our concern has been to introduce the general formalism of evolutionary genetics dynamics under fitness, in both the deterministic and stochastic setups, and chiefly in discrete-time. In the stochastic version of the problem, both the Wright-Fisher and the Moran models were considered. In the process, we revisited the various facets of the famous Fisher theorem of natural selection in both the deterministic and stochastic formulations. For the sake of simplicity of the exposition, we limited ourselves to a simplified one-parameter model where the sole selection parameter is unknown. Using these preliminary results and facts, we discussed the estimation problems of the selection parameter based on a single-generation frequency distribution shift under both deterministic and stochastic evolutionary dynamics. To the best of the author's knowledge, this particular way to address the estimation problem is new. It was stressed that in our models, there were no mutation effects included. We plan to include these effects in a forthcoming work. When mutations are present, the situation changes drastically. Firstly, in the deterministic formulation, the replicator dynamics combining fitness and mutations no longer is gradient-like in general. Still, an internal equilibrium point can exist were fitnesses to be multiplicative or would the mutation rates satisfy a house of cards condition [7]. Secondly, in the stochastic formulation, the Markov chains under study (either Wright-Fisher or Moran) are ergodic, now with an invariant measure which is independent of the initial condition. This also changes the picture drastically.

Acknowledgment. To a large extent, this work was triggered by the talk given by Professor Warren J. Ewens at the CIRM meeting in Marseilles, on May, 29, 2009; although the topics dealt here with are slightly different in spirit. Therefore, it owes much, if not all, to him. I take this opportunity to thank the organizers, Etienne Pardoux and Amaury Lambert, on behalf of the ANR MAEV headed by Sylvie Méléard, for giving me the chance to attend this Conference.

References

[1] Castilloux, A-M.; Lessard, S. The fundamental theorem of natural selection in Ewens' sense. Theor. Pop. Biol., 48, 306-315, 1995.

[2] Crow, J. F. Perspective: Here's to Fisher, additive genetic variance, and the fundamental theorem of natural selection. Evolution, 56(7), 1313-1316, 2002.

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

[4] Frank, S. A. The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis. Evolution, 51(6), 1712-1729, 1997.

[5] Frank, S. A. Natural selection maximizes Fisher information. J. of Evol. Biol., 22(2), 231-244, 2008.

[6] Kingman, J. F. C. A mathematical problem in population genetics. Proc. Cambridge Philos. Soc. 57, 574–582, 1961.

[7] 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

[8] Maruyama, T. *Stochastic problems in population genetics.* Lecture Notes in Biomathematics, 17. Springer-Verlag, Berlin-New York, 1977.

[9] Sella, G. An exact steady state solution of Fisher's geometric model and other models. Theoretical Population Biology, 75(1), 30-34, 2009.

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

Laboratoire de Physique Théorique et Modélisation, CNRS-UMR 8089 et Université de Cergy-Pontoise, 2 Avenue Adolphe Chauvin, F-95302, Cergy-Pontoise, France, Thierry.Huillet@u-cergy.fr