A General Approximation for the Dynamics of Quantitative Traits

Katarína Boďová,[1] Gašper Tkačik, and Nicholas H. Barton
Institute of Science and Technology Austria (IST Austria), Klosterneuburg A-3400, Austria

ABSTRACT Selection, mutation, and random drift affect the dynamics of allele frequencies and consequently of quantitative traits. While the macroscopic dynamics of quantitative traits can be measured, the underlying allele frequencies are typically unobserved. Can we understand how the macroscopic observables evolve without following these microscopic processes? This problem has been studied previously by analogy with statistical mechanics: the allele frequency distribution at each time point is approximated by the stationary form, which maximizes entropy. We explore the limitations of this method when mutation is small ($4N\mu < 1$) so that populations are typically close to fixation, and we extend the theory in this regime to account for changes in mutation strength. We consider a single diallelic locus either under directional selection or with overdominance and then generalize to multiple unlinked biallelic loci with unequal effects. We find that the maximum-entropy approximation is remarkably accurate, even when mutation and selection change rapidly.

KEYWORDS quantitative genetics; maximum entropy; diffusion approximation, quasi-stationarity

MOST traits of interest have a complex genetic basis depending on very many loci. Quantitative genetics gives a sophisticated statistical description of the components of trait variance that can predict the immediate change due to selection. The present abundance of genetic markers allows us to find some of the loci that affect traits, but such QTL typically account for only a small fraction of the genetic variance (Hill and Kirkpatrick 2010; Yang *et al.* 2010). While we may be able to predict breeding values and estimate the distribution of effects, it does not seem possible—even in principle—to identify the individual alleles responsible for the bulk of heritable variance. Thus, we cannot hope to predict the evolution of quantitative traits by using a direct population genetics approach based on the frequencies of each individual allele.

Here we develop a general method that allows us to closely approximate the evolution of quantitative traits knowing only the distribution of allelic effects and mutation rates but without requiring knowledge of individual allele frequencies. This can be seen as an extension of the classical infinitesimal model to include arbitrary gene interactions and the effects of selection, mutation, and drift on the genetic variance. It also can be viewed as a generalized version of the quasi-steady-state assumption (QSSA) that is often made in dynamical reaction systems (Segel and Slemrod 1989; Goeke and Walcher 2013) to noisy systems described by partial differential equations (PDEs), where the dynamics are approximated using a quasi-stationary distribution assumption (QSDA); here we use the maximum-entropy (MaxEnt) principle to define that distribution.

In physics, the MaxEnt principle has a long history, starting with the seminal work of Jaynes (1957), who interpreted the Boltzmann distribution of statistical physics as the most random distribution subject to a constraint on fixed average energy. In the recent decade, there has been a resurgence of interest in MaxEnt, especially when applied to biophysics problems ranging from the statistics of neural spiking (Schneidman *et al.* 2006; Tkačik *et al.* 2014), bird flocking (Bialek *et al.* 2012), protein structure (Weigt *et al.* 2009), and immunology (Mora *et al.* 2010). These approaches have been generalized to describe temporal dynamics of high-dimensional systems, known collectively as *maximum-caliber* or *dynamical/ kinetic maximum-entropy* models (Pressé *et al.* 2013), where the entropy of distribution over temporal paths is maximized given constraints on dynamical variables. Surprisingly,

however, MaxEnt distributions have not been used widely as a variational ansatz for cases discussed in this work, where the evolution equation for the distribution might be known *a priori*, *e.g.*, as is the case with the diffusion approximation in population genetics. In such a case, the approach differs from the maximum-caliber methods because it involves a combination of a static MaxEnt ansatz for the stationary microscopic distribution with a quasi-stationary assumption in the diffusion/ Fokker-Planck equation, which together extend the static MaxEnt inference to a dynamical approximation.

Prügel-Bennett and Shapiro (1997) and Rattray and Shapiro (2001) introduced the MaxEnt approximation to the dynamics of polygenic systems, predicting the cumulants of the trait distribution under mutation, selection, and drift; their main motivation was to understand evolutionary algorithms rather than natural populations. The same method was described independently in physics (Plastino *et al.* 1997) and used to approximate cosmic-ray fluxes (Hick and Stevens 1987).

However, neither of the two applications of MaxEnt was taken up in their respective fields. Barton and de Vladar (2009) modified Prügel-Bennet, Rattray, and Shapiro's method so that it could be justified from population genetics principles. With this modification, it gives the stationary distribution exactly and is accurate in the limit of slowly changing conditions. Numerical calculations showed that it is remarkably accurate, even when selection or mutation changes abruptly. However, the method is fully valid only when mutation is stronger than drift ($4N\mu > 1$), so populations are almost never fixed for one or other allele. Yet, in nature, mutation is typically weaker than drift, so most sites are fixed; in this case, Barton and de Vladar's approximation applies only in cases where mutation or population size does not change with time.

We begin by summarizing the stationary MaxEnt approximation and its extension to nonstationary problems. We then extend the MaxEnt approximation so that it applies over the full range of mutation rates and test the accuracy of this approximation for directional selection and for balancing selection that favors heterozygotes. This extension is a combination of the continuous approach of Barton and de Vladar (2009) and a special treatment of the dynamics at the boundaries. A similar approach, where the boundaries have to be treated differently, has been used in the semidiscrete, semicontinuous methods studied in reaction-diffusion systems (Flegg *et al.* 2011; Robinson *et al.* 2014) and also in traveling fitness waves, where the fluctuations can be introduced to the model using a cutoff function (Tsimring *et al.* 1996; Hallatschek 2011). We initially consider the distribution of allele frequencies at a single locus and then extend consideration to multiple loci with a distribution of effects. Throughout, we assume that recombination is fast relative to other processes so that the population is in linkage equilibrium and can be described by its allele frequencies.

Dynamics of Allele Frequencies

The dynamics of allele frequencies $\mathbf{p} = (p_1, \ldots, p_L)$ (for biallelic loci) can be described by a diffusion process using a deterministic forward Kolmogorov equation (*i.e.*, the Fokker-Planck or diffusion equation). The evolution of the joint probability density $\psi(\mathbf{p}, t)$ of allele frequencies satisfies

$$\frac{\partial \psi}{\partial t} = -\sum_{i=1}^{L} \frac{\partial}{\partial p_i}[\mathcal{M}_i(p_i)\psi] + \frac{1}{2}\sum_{i=1}^{L}\frac{\partial^2}{\partial p_i^2}[\mathcal{V}_i(p_i)\psi] \quad (1)$$

where the number of loci that contribute to the trait is L.

The second term of (1) equals $\mathcal{V}_i(p_i) = p_i q_i/2N$ and captures stochasticity of the allele frequencies arising from random sampling, *i.e.*, the random drift. While in the case of linkage disequilibrium this term would contain a double summation, reflecting correlations between loci, the off-diagonal terms vanish at linkage equilibrium. (The factors of 2 in the brackets of (1) arise because we assume a diploid population of N individuals; the corresponding haploid model would be the same, apart from these factors.)

The first term of (1) captures deterministic changes of allele frequencies. We consider

$$\mathcal{M}_i(p_i) = (\beta\gamma_i + h\eta_i)p_i q_i\left(1 - \frac{2h\eta_i}{\beta\gamma_i + h\eta_i}p_i\right) + \mu q_i - \nu p_i \quad (2)$$

where β is the strength of directional selection, h denotes a higher-order correction that captures dominance, μ and ν are the forward and backward mutation rates, and γ_i and η_i are the additive effects of the ith locus on the traits under selection. $\mathcal{M}_i(p_i)$ can be written in a potential form

$$\mathcal{M}_i(p_i) = \frac{p_i q_i}{2}\frac{\partial(\boldsymbol{\alpha} \cdot \mathbf{A})}{\partial p_i}$$

where the potential $\boldsymbol{\alpha} \cdot \mathbf{A}$, obtained by inverting this relationship, reflects effects of selection and mutation:

$$\boldsymbol{\alpha} \cdot \mathbf{A} = \sum_j \alpha_j A_j = \log\overline{W} + \mu U + \nu V$$
$$= \beta\bar{z} + hH + 2\mu\sum_{i=1}^{L}\log p_i + 2\nu\sum_{i=1}^{L}\log q_i \quad (3)$$

with quantities $\bar{z}$, U, V, and H defined as

$$\bar{z} = \sum_{i=1}^{L}\gamma_i(p_i - q_i), \quad H = \sum_{i=1}^{L}2\eta_i p_i q_i \quad (4)$$

$$U = \sum_{i=1}^{L}2\log p_i, \quad V = \sum_{i=1}^{L}2\log q_i \quad (5)$$

where $q_i = 1 - p_i$, and γ_i and η_i are the effects of loci on the traits $\bar{z}$ and H, respectively.

Selection, captured by the mean fitness $\overline{W}$, may be an arbitrary function of the distribution of quantitative traits, which, in turn, may be an arbitrary function of the n allele frequencies. The log mean fitness may be further decomposed into a sum in which the $\boldsymbol{\alpha}_{\overline{W}}$ represent various

selection coefficients and $A_{\overline{W}}$ various selected traits. In the case of dominance, the terms associated with selection are $\log \overline{W} = \beta\overline{z} + hH = \boldsymbol{\alpha}_{\overline{W}} \cdot \mathbf{A}_{\overline{W}}$, where $\boldsymbol{\alpha}_{\overline{W}} = (\beta, h)$ and $\mathbf{A}_{\overline{W}} = (\overline{z}, H)$. Here we assume a weak selection $\beta \ll 1$ such that $\overline{\log W} \approx \log \overline{W}$.

The deterministic effects on allele frequency can be summarized into a vector of coefficients $\boldsymbol{\alpha}$ and a vector of complementary quantities $\mathbf{A}$. We study directional selection and dominance with nonsymmetrical mutation and define $\boldsymbol{\alpha}$ and $\mathbf{A}$ as

$$\boldsymbol{\alpha} = (\beta, h, \mu, \nu), \quad \mathbf{A} = (\overline{z}, H, U, V) \qquad (6)$$

In the sections that follow, we show that $\mathbf{A}$ and $\boldsymbol{\alpha}$ can be understood as constraints and corresponding Lagrange multipliers, respectively, of a particular variational problem.

For $h = 0$ and $\mu = \nu$, this represents the simplest scenario of directional selection with symmetrical mutation. Directional selection of strength β acts on a trait $\overline{z}$, assumed to be additive, while selection of strength h acts on heterozygosity H. In this work, we consider unequal effects γ_i on the trait but equal effects $\eta_i = 1$ on H. This can be easily extended to distribution of effects η_i. A wide variety of other models can be treated in the same way. For example, de Vladar and Barton (2011) studied stabilizing selection on an additive trait.

This diffusion process is known to be an accurate continuous-time approximation to a wide range of specific population genetics models (Kimura 1955a; Ewens 2012); moreover, it corresponds directly to the coalescent process that describes the ancestry of samples taken from the population (Wakeley 2008). In order to represent the population in terms of allele frequencies, we must assume that linkage disequilibria are negligible, which will be accurate if recombination is sufficiently fast. For simplicity, we also assume two alleles per locus.

The stationary distribution of (1) has the form (Wright 1931; Kimura 1955a)

$$\begin{aligned}
\overline{\psi}(\mathbf{p}) &= \frac{1}{\mathbb{Z}} \exp(2N\boldsymbol{\alpha} \cdot \mathbf{A}) \prod_{i=1}^{L} (p_i q_i)^{-1} \\
&= \frac{1}{\mathbb{Z}} \overline{W}^{2N} \prod_{i=1}^{L} p_i^{4N\mu-1} q_i^{4N\nu-1}
\end{aligned} \qquad (7)$$

where $\mathbb{Z}$ is a normalization constant, also called a *partition function*. This distribution falls to zero at the boundaries ($p = 0, 1$) provided that $4N\mu > 1$ and $4N\nu > 1$. However, when mutation rates fall below this threshold, the distribution develops singularities at boundaries (if $4N\mu$ or $4N\nu$ is small, the singularity occurs at $p = 0$ or $p = 1$, respectively), even though the density function is still integrable.

MaxEnt in Equilibrium Quantitative Genetics

The stationary distribution (7) can be derived from a variational MaxEnt principle. The key assumption is that selection and mutation act only through a set of observable quantities, which can be arbitrary functions of allele frequencies $\mathbf{A} = \{A_i(p)\}$; the strength of selection and mutation is given by the corresponding set of $\boldsymbol{\alpha} = \{\alpha_i\}$. Together these define the potential function $\boldsymbol{\alpha} \cdot \mathbf{A}$.

We can define an entropy, relative to a reference measure $\phi(\mathbf{p})$, as

$$S_H[\psi] = - \int\limits_{[0,1]^L} \psi(\mathbf{p}, t) \log\left[\frac{\psi(\mathbf{p}, t)}{\phi(\mathbf{p})}\right] d\mathbf{p} \qquad (8)$$

which has a unique maximum $\psi(\mathbf{p}, t) = \phi(\mathbf{p})$ at the reference distribution; the entropy is (minus) the Kullback-Leibler divergence from $\phi(\mathbf{p})$. The key choice is to set the reference distribution as the neutral distribution of allele frequencies in the absence of mutation or selection:

$$\phi(\mathbf{p}) = \prod_{i=1}^{L} (p_i q_i)^{-1} \qquad (9)$$

Note that $\phi(\mathbf{p})$ is not integrable, but it does provide the neutral probability distribution given that the allele is not fixed and yields the stationary distribution under mutation, selection, and drift when we maximize S_H subject to a normalization constraint and constraints $\langle\mathbf{A}\rangle = \langle\mathbf{A}\rangle_{\text{obs}}$. The latter condition enforces a constraint on the ensemble averages $\langle\mathbf{A}(\mathbf{p})\rangle = \int \mathbf{A}(\mathbf{p})\psi(\mathbf{p})d\mathbf{p}$. These macroscopic quantities represent information that, in principle, could be observed. We refer to $\langle\mathbf{A}(\mathbf{p})\rangle$ as observables, even though this does not necessarily mean that their values over time are known. The constrained maximization of entropy is solved by a method of Lagrange multipliers; for details, see Appendix A. For the example in (6), the constraints include

1. $\int_{[0,1]^L} \psi(\mathbf{p}) d\mathbf{p} = 1$ – normalization constraint,
2. $\langle\overline{z}\rangle = \langle\overline{z}\rangle_{\text{obs}}$ with Lagrange multiplier $2N\beta$,
3. $\langle H\rangle = \langle H\rangle_{\text{obs}}$ with Lagrange multiplier $2Nh$,
4. $\langle U\rangle = U_{\text{obs}}$ with Lagrange multiplier $2N\mu$, and
5. $\langle V\rangle = V_{\text{obs}}$ with Lagrange multiplier $2N\nu$.

The normalization condition sets the total probability of the allele frequency distribution to 1 and introduces the partition function $\mathbb{Z}$ as a constant multiplier in (7).

This variational principle recovers the stationary distribution of the diffusion equation (1); the $2N$ times the mutation rates and selection coefficients can be thought of as the Lagrange multipliers. Iwasa (1988) introduced the same entropy measure but used it in a slightly different way: he showed that the sum of the potential function and $S_H/2N$ define a free fitness that increases over time, just as in thermodynamics the free energy increases over time. Further connections to thermodynamics have been studied by Sella and Hirsh (2005) for a special case of very small mutations, where most of the alleles are fixed, and in Barton and Coe (2009), including the novel entropy term (analogous to our U and V) that involves the effects of mutation.

Note that $\mathbf{A}$ and $\boldsymbol{\alpha}$ have a specific meaning both in quantitative genetics and in statistical physics. In quanti-
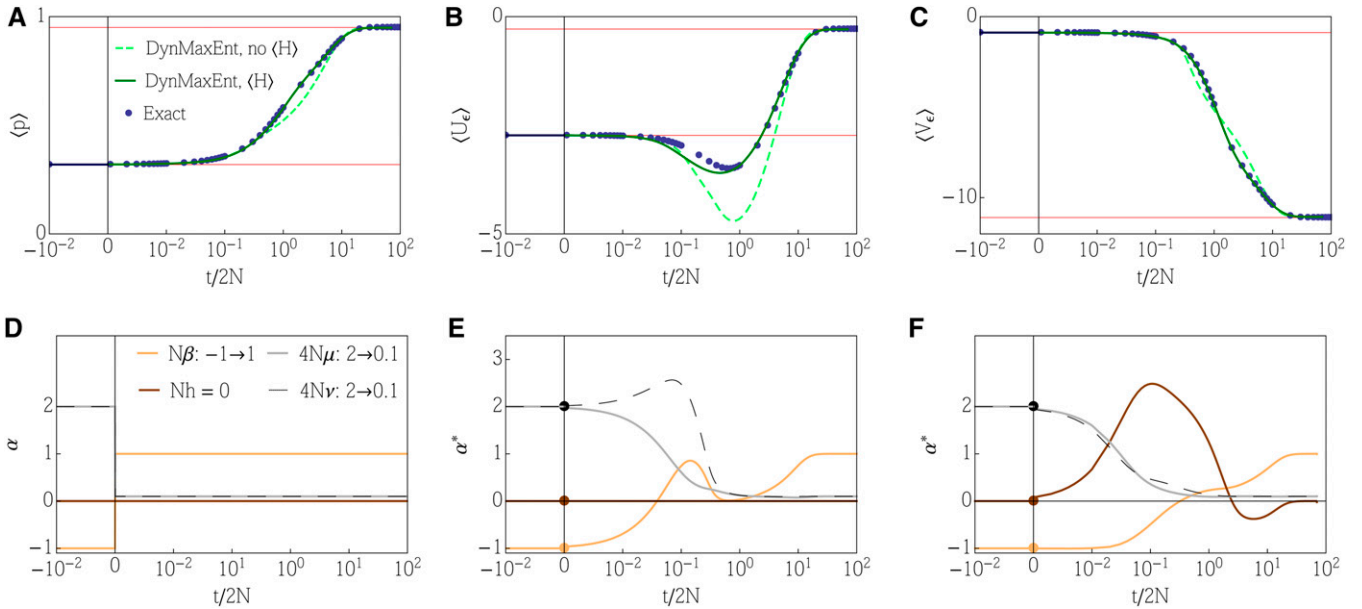
Figure 1 Example 1: details of the general DynMaxEnt method. The response of the observed quantities: (A–C) Observables $\langle p \rangle$, $\langle U_\varepsilon \rangle$, and $\langle V_\varepsilon \rangle$ obtained by numerically solving the diffusion equation using a discretization of space and time where the explicit transition matrix is known (provided in a Mathematica notebook in File S1) in blue dots and the general DynMaxEnt approximations in green color. We used $2N = 1000$ and $\varepsilon = 1/2N$. The dashed green line shows approximation with the three shown observables and the solid green line uses an additional observable $\langle pq \rangle$. (D) Changes in evolutionary forces that draw the system out of equilibrium include a rapid decrease in mutation complemented by a change in selection strength. (E) Dynamics of Lagrange multipliers in the general DynMaxEnt method when three constraints are employed: $\langle p \rangle$, $\langle U_\varepsilon \rangle$, and $\langle V_\varepsilon \rangle$. (F) Dynamics of Lagrange multipliers in the general DynMaxEnt method when four constraints are employed, including $\langle pq \rangle$.

genetics, **A** characterizes properties of a quantitative trait whose means can be observed ($\langle \mathbf{A} \rangle$) and that evolves in response to the evolutionary forces $\boldsymbol{\alpha}$. In statistical physics, **A** and $\boldsymbol{\alpha}$ represent conjugate pairs of thermodynamic variables, which can be interpreted as the constraints and Lagrange multipliers in the variational MaxEnt problem, commonly encountered when microscopic states of the system are unobserved but its macroscopic features are known.

Note that there is some flexibility in the choice of the reference distribution $\phi(\mathbf{p})$, where different choices may lead to the same stationary distribution. For instance, one may take the neutral distribution that involves mutation terms $\phi(\mathbf{p}) = \prod_k p_i^{4N\mu - 1} q_i^{4N\nu - 1}$ while omitting the constraints on U and V and assuming that μ and ν are functions of time. In this way, the reference distribution would be normalizable. Because these approaches are equivalent, one can view the constraints on U and V as conditions that regularize the allele frequency distribution.

Dynamic MaxEnt Approximation

Our aim is to approximate the dynamics of a high-dimensional system by a small number of variables, which include the quantities that determine fitness. We approximate the real distribution of allele frequencies by the stationary distribution obtained by the MaxEnt method with a small number of constraints and use it as an ansatz in the diffusion equation. This leads to effective dynamical forces $\boldsymbol{\alpha}^*$ that yield the correct dynamics for the observed quantities. The assumption

that the population is perturbed only through the forces $\boldsymbol{\alpha}$ is crucial to the success of our approximation. If we could manipulate individual allele frequencies in an arbitrary way, then the long-term evolution would become essentially unpredictable: alleles that are initially rare could increase to cause arbitrary changes as they eventually rose to appreciable frequency (Barton and de Vladar 2009) (Figure 1). The overall strategy of the dynamical MaxEnt (DynMaxEnt) approximation is summarized in Table 1, while the terminology from statistical physics and quantitative genetics is provided in Table 2. Various approximate methods, discussed in our work, are summarized in Table 3.

We first describe the continuous DynMaxEnt method, proposed in Barton and de Vladar (2009), which requires a sufficient number of mutations in every generation. However, a discrete approximation, also used in Barton and de Vladar (2009) and described in Appendix C, is applicable when the mutation rate is small and selection is limited. The dynamics are then formulated in terms of fixed classes of alleles. However, the discrete approximation is not accurate unless the mutation rate is very small, and even then it has a limitation when $N\beta \gg 1$ (Appendix C). Similarly, the continuous method fails for $4N\mu < 1$ (Appendix D). We compare the performance of these methods in Appendix G and find that while the discrete method applies to a very small mutation rate and the continuous method to a large mutation rate, the intermediate regime is not captured by either of them. This is similar to the result of Mustonen and Lässig (2007, 2008), who studied fitness waves in the problem of fluctuating

Table 1 Summary of the DynMaxEnt approach in four steps

Step 1	Formulate dynamics, as in (1), for the probability distribution of the state variables $\psi(p)$.
Step 2	Obtain the stationary distribution ψ and write it in an exponential (log-linear) form $\psi(p) \propto \phi(p)\exp(2N\boldsymbol{\alpha} \cdot \mathbf{A})$ in terms of observables $\langle \mathbf{A} \rangle$ and constant forces $\boldsymbol{\alpha}$.
Step 3	Represent $\psi(p)$ as a solution of a variational MaxEnt problem with reference distribution $\phi(p)$, constraints on $\langle \mathbf{A} \rangle$, and Lagrange multipliers $\boldsymbol{\alpha}$ (nonunique).
Step 4	Use a quasi-stationarity assumption to approximate the dynamics of observables using the stationary distribution where the coefficients $\boldsymbol{\alpha}$ are allowed to change over time to match the correct dynamics of observables. This criterion leads to a reduced dynamical system for the effective coefficients $\boldsymbol{\alpha}^*$.

See File S1.

selection and introduced a novel approximation that is accurate for small selection timescales. Figure 2 in Mustonen and Lässig (2008) also shows an intermediate regime in which neither the diffusion theory nor the novel approximation is accurate. In this work, we present a *general DynMaxEnt* approximation that is applicable in all regimes. This approximation is compared to the numerical solution of the diffusion equation. Instead of using individual-based simulations, which are computationally demanding, we consider allele frequencies to hold biologically relevant values $p_k = k/2N$, for $k = 0, \ldots, N$, and forward iterate an explicit transition matrix, consistent with the diffusion equation. This approach, provided in a Mathematica notebook (Supplemental Material, File S1), is also feasible when a moderate number of loci have different effects on the trait of interest.

Continuous DynMaxEnt approximation

Any set of forces $\boldsymbol{\alpha}$ will cause the population to evolve to a stationary distribution $(\phi/\mathbb{Z})\exp(2N\boldsymbol{\alpha} \cdot \mathbf{A})$; this is the distribution that maximizes entropy subject to constraints on $\langle \mathbf{A} \rangle$, the $2N\boldsymbol{\alpha}$ being the Lagrange multipliers. Now suppose that the forces change abruptly, from $\boldsymbol{\alpha}_0$ to $\boldsymbol{\alpha}_1$, and no further information about the system is provided. The expected observables $\langle \mathbf{A} \rangle$ will evolve toward the new stationary distribution. At any time, there will be a set of forces $\boldsymbol{\alpha}^*$ that would produce the current $\langle \mathbf{A} \rangle$ if the population were stationary; we expect that the $\boldsymbol{\alpha}^*$ will evolve from $\boldsymbol{\alpha}_0$ to $\boldsymbol{\alpha}_1$ as the population evolves from one stationary state to the other. Thus, we can describe the dynamics either by the change in $\langle \mathbf{A} \rangle$ or, equivalently, by the change in $\boldsymbol{\alpha}^*$.

Under the diffusion approximation, the expectations change as

$$\frac{\partial \langle A_i \rangle}{\partial t} = \sum_j B_{i,j}\alpha_j + \frac{1}{2N}V_i \qquad (10)$$

where

$$B_{i,j} = \left\langle \sum_k \frac{\partial A_i}{\partial p_k}\frac{p_k q_k}{2}\frac{\partial A_j}{\partial p_k} \right\rangle, \quad V_i = \left\langle \sum_k \frac{\partial^2 A_i}{\partial p_k^2}\frac{p_k q_k}{2} \right\rangle \qquad (11)$$

[Equations 13 and 14 of Barton and de Vladar (2009); note that in their Equation 13 the expectation should be taken over the whole equation, not inside the derivatives as typed].

The expectations that appear on the left-hand side of (10) are not the same as the ones on the right; therefore, the system is not closed. We now introduce the continuous DynMaxEnt approximation, namely, that $B_{i,j}$ and V_i in the dynamically changing system are approximated by the values that they would have at the corresponding stationary state that generates the actual $\langle \mathbf{A} \rangle$; the stationary distribution coincides with the MaxEnt distribution. If the population were at a stationary state under the forces $\boldsymbol{\alpha}^*$, chosen to produce the current expectations $\langle \mathbf{A} \rangle$, then there would be no change:

$$\frac{\partial \langle A_i \rangle}{\partial t} = 0 = \sum_j B_{i,j}^*\alpha_j^* + \frac{1}{2N}V_i^* \qquad (12)$$

where the asterisks denote values at the stationary distribution. The approximation has form $(1/2N)V_i^* \approx -\sum_j B_{i,j}^*\alpha_j^*$, which gives Equation 15 of Barton and de Vladar (2009) as

$$\frac{\partial \langle A_i \rangle}{\partial t} = \sum_j B_{i,j}^*\left(\alpha_j - \alpha_j^*\right) \qquad (13)$$

It may be more convenient to follow the rates of change in $\boldsymbol{\alpha}^*$, which can be written in terms of the covariance of fluctuations in the $\mathbf{A}$. Using matrix notation,

$$\frac{\partial \boldsymbol{\alpha}^*}{\partial t} = \frac{1}{2N}\mathbf{C}^{*-1}\mathbf{B}^* \cdot (\boldsymbol{\alpha} - \boldsymbol{\alpha}^*) \qquad (14)$$

where

$$\mathbf{C}^* = \text{Cov}(A_i, A_j) = \left\langle \frac{\partial^2 \log \mathbb{Z}}{\partial \alpha_i \partial \alpha_j} \right\rangle \qquad (15)$$

with an initial condition $\boldsymbol{\alpha}^*(0) = \boldsymbol{\alpha}_0$ and $\boldsymbol{\alpha} = \boldsymbol{\alpha}_1$. The difference $|\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0|$ represents the change in evolutionary forces. Because the matrices $\mathbf{B}^*$ and $\mathbf{C}^*$ depend only on the effective forces $\boldsymbol{\alpha}^*$, as shown in Appendix B, the dynamical system for $\boldsymbol{\alpha}^*$ is closed. A detailed derivation of the DynMaxEnt method under more general conditions can be found in Appendix E.

Intuitively, one may assume quasi-stationarity in (10), provided that the evolutionary forces $\boldsymbol{\alpha}$ change on a slower timescale than the timescales of selection $(1/\beta)$, mutation $(1/\mu)$, and random drift $(2N)$. Then the adiabatic approximation in (14) should be accurate not only for the predicted observables but also for the microscopic distribution. However, we will show that even when the evolutionary forces change

Table 2 Table of key terms and constants

Symbol	Quantitative genetics/statistical physics
$\mathbf{A}$	Macroscopic observables
$\langle\mathbf{A}\rangle$	Constraints
$\boldsymbol{\alpha}$	Evolutionary forces/Lagrange multipliers
$\psi(p,t)$	Distribution of allele frequencies/Boltzmann distribution
$\phi(p)$	Neutral distribution (no mutation)/reference distribution
S_{H}	Relative entropy, negative Kullback-Leibler divergence
$\overline{W}$	Mean fitness
$2N$	Population size in a diploid population
L	Number of loci contributing to the trait
$\mathbf{B}$	Additive genetic covariance matrix
$\mathbf{C}$	Covariance in fluctuations/susceptibility matrix
$J[\psi,p]$	Flux of the probability mass at frequency p

abruptly, *i.e.*, when $|\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0|$ is not small, the approximation remains accurate—even though there is no guarantee that the inferred microscopic distribution agrees with the correct one.

The matrix $\mathbf{B}$ can be seen as a generalization of the additive genetic covariance matrix, where the $\partial A_{\mathrm{i}}/\partial p_{\mathrm{k}}$ corresponds to (twice) the marginal effects of the kth allele on A_{i}. The MaxEnt approximation consists of assuming that this matrix is approximately what one would obtain at equilibrium with the current $\langle\mathbf{A}\rangle$. Thus, (13) is an extension to the "breeder's equation" (Lynch and Walsh 1997), which allows for quantities $\langle\mathbf{A}\rangle$ that can be any function of allele frequencies—not just trait means—and that allows for random fluctuations, mutational bias, and nonadditive selection.

The DynMaxEnt method can be contrasted with the maximum-caliber method, reviewed in Pressé *et al.* (2013). DynMaxEnt uses static observables to infer the correct stationary allele frequency distribution but allows the Lagrange multipliers to change over time in accordance with the known diffusion equation, ensuring that the observables are correct at all times. However, the maximum-caliber method uses constraints on temporal characteristics to arrive at a distribution over the allele frequency paths with constant values of Lagrange multipliers. DynMaxEnt is suitable for our problem because it only assumes knowledge of initial and changed evolutionary parameters and no further information on the properties of the allele frequency paths.

General DynMaxEnt approximation

When numbers of mutations are small (*i.e.*, $4N\mu < 1$ and $4N\nu < 1$), we face a problem of diverging components in the continuous DynMaxEnt approximation as a result of a U-shaped allele frequency distribution (for simplicity, we will consider a single locus). If $\mathbf{A} = (\overline{z}, H, U, V)$, this divergence is caused by diagonal elements of matrix $\mathbf{B}$ that correspond to U and V, in particular, $B_{3,3} = \langle pq/2 \cdot [\partial_p U(p)]^2\rangle = \langle 2q/p\rangle$ for $4N\mu < 1$ and $B_{4,4} = \langle pq/2 \cdot [\partial_p V(p)]^2\rangle = \langle 2p/q\rangle$ for $4N\nu < 1$ (see Appendices B and D for more detail). Therefore, the continuous DynMaxEnt approximation fails completely in a regime of dynamic selection and mutation when $4N\mu < 1$ or $4N\nu < 1$ simply because the right-hand side of the dynamical system (14) is ill defined. The breakdown of the continuous

DynMaxEnt method, when the number of mutations are small (*i.e.*, for small populations), is not a numerical problem but a fundamental limitation of the method itself. However, it can be avoided by considering a modified diffusion problem that does not aim to resolve all details of the allele frequency distribution close to the fixation and loss but instead agrees with the original problem in terms of the probability that the allele frequency is extreme.

We define the boundary layers as $[0, \varepsilon]$ and $(1 - \varepsilon, 1]$ for $\varepsilon \ll 1$ but finite. The value of the truncation parameter ε is discussed later, but typically, $\varepsilon \sim 1/N$. We then replace the original diffusion equation with solution $\psi(p,t)$ by a new system of PDEs with solution $\psi_\varepsilon(p,t)$ that agrees with the true dynamics in the following properties:

1. The stationary distribution in the bulk is the same for both problems: $\psi_\varepsilon(p, t = \infty) = \psi(p, t = \infty)$, for $p \in [\varepsilon, 1 - \varepsilon]$.
2. The stationary probabilities of extreme allele frequencies are the same for both problems:

$$P[p < \varepsilon] = \int_0^\varepsilon \psi(p, t = \infty)dp = \int_0^\varepsilon \psi_\varepsilon(p, t = \infty)dp$$

and

$$P[p > 1 - \varepsilon] = \int_{1-\varepsilon}^1 \psi(p, t = \infty)dp = \int_{1-\varepsilon}^1 \psi_\varepsilon(p, t = \infty)dp$$

to the lowest order in ε.

3. As the truncation parameter goes to 0, the problem converges to the original diffusion equation; *i.e.*, it develops singularities at the boundaries:

$$\lim_{\varepsilon \to 0} \psi_\varepsilon(p,t) = \psi(p,t).$$

Replacing the original diffusion equation with a set of coupled diffusion equations in different regions of the state space captures an important characteristic of the problem: the presence of multiple timescales in the allele frequency dynamics. When the system is perturbed from stationarity by a change in the Lagrange multipliers (selection, mutation, heterozygosity), *e.g.*, by a dramatic drop in number of mutations, the correct distribution very quickly adjusts to the form $p^{4N\mu-1}q^{4N\nu-1}$ at the boundaries. Only then does the mass in the interior slowly transfer to the vicinity of the fixed states to converge to the new stationary distribution. This results in a very fast dynamics of $\langle U\rangle$ and $\langle V\rangle$ and a considerably slower dynamics of the trait mean and the mean heterozygosity.

DynMaxEnt can capture these multiscale features only if it can incorporate a low and changing rate of mutation. The quick initial adjustment of the mutation rates is then naturally followed by a slow dynamics of the trait mean and heterozygosity because the speed of the transfer of the mass between the fixed states is limited by the infrequent rate of mutation. Similarly to

Table 3 Overview of the approximations

Abbreviation	General approach
MaxEnt	Stationary microscopic distribution derived from macroscopic observables.
DynMaxEnt	Any approximation for the dynamics of observables based on a combination of MaxEnt variational ansatz, quasi-stationarity assumption, and a dynamical (*e.g.*, diffusion) equation for the probability distribution.

Abbreviation	Method in quantitative genetics
Continuous DynMaxEnt	Refers to the method of Barton and de Vladar (2009), where the dynamics are captured by diffusion equation (1).
General DynMaxEnt	The approximation introduced here, which generalizes the continuous DynMaxEnt method. This new approximation distinguishes between the bulk mass that behaves as in the continuous problem and the coupled boundary masses that behave as discrete quantities.

other observables, the boundary masses also can be expressed as means of functions of allele frequency and thus treated as additional observables. The appropriate function is a characteristic function $\chi_{[a,b]}$ where

$$\chi_{[a,b]} = \begin{cases} 1 & p \in [a,b] \\ 0 & \text{else} \end{cases} \tag{16}$$

Splitting the problem domain augments the degrees of freedom by $\langle \chi_{[0,\varepsilon]} \rangle$ and $\langle \chi_{[1-\varepsilon,1]} \rangle$, which control the boundary dynamics independently of the bulk dynamics and the corresponding Lagrange multipliers $2N\kappa$ and $2N\rho$. The values of κ and ρ will be determined later. The general DynMaxEnt method, derived in Appendix E and summarized next, provides a way to couple the boundary dynamics with the bulk dynamics to account for the multiscale features and resolve the technical problems of the continuous method while keeping the same number of degrees of freedom as the continuous DynMaxEnt.

We first split the allele frequency domain into the bulk part and the boundary part and define the diffusion equation separately in the three regions in Appendix E. We couple the equations by a boundary flux that is consistent with the original diffusion equation, thus leading to the same probability mass at the boundaries as in the continuous DynMaxEnt method. The stationary distribution of the problem (E4) has the form

$$\overline{\psi}_\varepsilon(p) = \frac{1}{\mathbb{Z}} \overline{W}^{2N} \begin{cases} e^{2N\kappa} \varepsilon^{4N\mu-1} q^{4N\nu-1} & \text{if } p < \varepsilon \\ p^{4N\mu-1} q^{4N\nu-1} & \text{otherwise} \\ e^{2N\rho} p^{4N\mu-1} \varepsilon^{4N\nu-1} & \text{if } p > 1-\varepsilon \end{cases} \tag{17}$$

with $\mathbb{Z} = \int \psi_\varepsilon(p) dp$, where $\mathbb{Z}$ is the normalization constant, and the relative masses in the three regions are determined by constants κ and ρ. The stationary distribution (17) is not generally continuous at $p = \varepsilon, 1-\varepsilon$. However, it can still be obtained by maximizing a relative entropy with a bounded base distribution

$$\phi_\varepsilon(p) = \frac{1}{\mathbb{Z}^0} \begin{cases} (\varepsilon q)^{-1} & \text{if } p < \varepsilon \\ (pq)^{-1} & \text{otherwise} \\ (p\varepsilon)^{-1} & \text{if } p > 1-\varepsilon \end{cases} \qquad \mathbb{Z}^0 = \int_0^1 \phi_\varepsilon(p) dp \tag{18}$$

complemented by the following constraints on $\langle \mathbf{A} \rangle$ and Lagrange multipliers $2N\boldsymbol{\alpha}$:

$$\mathbf{A} = (\log \overline{W}, U_\varepsilon, V_\varepsilon) \tag{19}$$

$$\boldsymbol{\alpha} = (\alpha_{\overline{W}}, \mu, \nu) \tag{20}$$

where instead of U and V that diverge at the boundaries, we take their truncation to $p \in [\varepsilon, 1-\varepsilon]$:

$$U_\varepsilon = \begin{cases} 2\log \varepsilon & p < \varepsilon \\ 2\log p & \text{else} \end{cases}, \quad V_\varepsilon = \begin{cases} 2\log q & p < 1-\varepsilon \\ 2\log \varepsilon & \text{else} \end{cases} \tag{21}$$

The two remaining parameters, κ and ρ, are matched to satisfy conditions 2 from the preceding list, leading to

$$\kappa = -\frac{1}{2N}\log(4N\mu) \quad \text{and} \quad \rho = -\frac{1}{2N}\log(4N\nu) \tag{22}$$

This relationship ensures that the approximate stationary distribution has the same proportion of the mass at the boundaries to the stationary solution of the original diffusion. Generalized to multiple loci, the MaxEnt distribution has the form

$$\overline{\psi}_\varepsilon(\mathbf{p}) = \frac{1}{\mathbb{Z}} \overline{W}^{2N} \prod_{i=1}^{L} \begin{cases} \frac{1}{4N\mu} \varepsilon^{4N\mu-1} q_i^{4N\nu-1} & \text{if } p_i < \varepsilon \\ p_i^{4N\mu-1} q_i^{4N\nu-1} & \text{otherwise} \\ \frac{1}{4N\nu} p_i^{4N\mu-1} \varepsilon^{4N\nu-1} & \text{if } p_i > 1-\varepsilon \end{cases} \tag{23}$$

with $\mathbb{Z}$ such that $1 = \int \psi_\varepsilon(p) dp$. We remark that the reference distribution $\phi_\varepsilon(p)$ is the stationary distribution in the absence of selection and mutation. Therefore, $\phi_\varepsilon(p) = \sigma_\varepsilon(p)^{-1}$, where σ_ε is the state-dependent diffusion coefficient:

$$\sigma_\varepsilon(p) = \begin{cases} \varepsilon q & p < \varepsilon \\ pq & \text{else} \\ p\varepsilon & p > 1-\varepsilon \end{cases} \tag{24}$$

The diffusion equation in the split domain can be used to derive a new DynMaxEnt approximation for arbitrary mutation strengths. This is done in a manner similar to the

continuous DynMaxEnt method using the ansatz (17) and a quasi-stationarity assumption. The derivation, detailed in Appendix E, leads to the same dynamics as the original method:

$$\frac{\partial \langle \mathbf{A} \rangle}{\partial t} = \mathbf{B}_\varepsilon (\boldsymbol{\alpha} - \boldsymbol{\alpha}^\star) \tag{25}$$

$$\frac{\partial \boldsymbol{\alpha}^\star}{\partial t} = \frac{1}{2N} \mathbf{C}_\varepsilon^{-1} \mathbf{B}_\varepsilon (\boldsymbol{\alpha} - \boldsymbol{\alpha}^\star) \tag{26}$$

where (26) is a closed dynamical system for $\boldsymbol{\alpha}^\star$. However, since constant prefactors in the stationary distribution (23) depend on mutation rates, the matrix $\mathbf{C}_\varepsilon$ contains additional terms:

$$\mathbf{B}_\varepsilon = \left\langle \frac{\sigma_\varepsilon(p)}{2} \frac{\partial A_j}{\partial p} \frac{\partial A_k}{\partial p} \right\rangle, \quad \mathbf{C}_\varepsilon = \mathrm{Cov}(A_j, A_k)$$

$$- \left[0_{\overline{W}} \quad \frac{1}{2N\mu} \mathrm{Cov}\left(\mathbf{A}, \chi_{[0,\varepsilon)} \right) \quad \frac{1}{2N\nu} \mathrm{Cov}\left(\mathbf{A}, \chi_{[1-\varepsilon,1)} \right) \right] \tag{27}$$

where the subtracted terms form a square matrix of the same dimension as $\mathrm{Cov}(A_j, A_k)$. Its leftmost columns, denoted by $0_{\overline{W}}$, represent a contribution to the terms $\mathrm{Cov}(\cdot, \mathbf{A}_{\overline{W}})$. Because the stationary distribution (23) depends on the selection-related Lagrange multipliers only through $\overline{W}^{2N} = e^{\boldsymbol{\alpha}_{\overline{W}} \cdot \mathbf{A}_{\overline{W}}}$, in a similar manner to the continuous DynMaxEnt method, all added terms are zero. In contrast, the presence of the mutation-related Lagrange multipliers in (23) as the scaling factors results in additional terms added to $\mathrm{Cov}(\cdot, U_\varepsilon)$ and $\mathrm{Cov}(\cdot, V_\varepsilon)$, forming the remaining two columns in (27). The components of the matrices $\mathbf{B}_\varepsilon$ and $\mathbf{C}_\varepsilon$ are functions of effective forces only, similar to the continuous DynMaxEnt approximation. The difference is that the expectations involve integrals through parts of the domain $[0, 1]$ (bulk part or boundary parts), and thus the terms have to be evaluated numerically.

Applications of the General DynMaxEnt Method

In the following, we show the performance of the general DynMaxEnt method in the parameter regimes when the continuous DynMaxEnt method is not applicable because the dynamics (14) become singular. These examples involve dynamic selection and mutation in the regime $4N\mu, 4N\nu < 1$. To illustrate the method step by step, we provide a pedagogical Mathematica notebook with the code for the general DynMaxEnt method (File S1).

Performance of the DynMaxEnt approximation for different scenarios is compared to the discrete and continuous approximations in Appendix G.

Example 1: Low and changing mutation

Here we consider a single locus under directional selection. As we show in Appendix E, this also corresponds to the case of multiple loci with equal effects. Despite its apparent simplicity,

the system still contains $2N$ degrees of freedom that capture the allele frequency distribution in a population of N diploid individuals. The general DynMaxEnt approximation, based on a stationary MaxEnt approximation, reduces the dynamics to a few degrees of freedom that correspond to the Lagrange multipliers $\boldsymbol{\alpha}^\star(t)$. For instance, when $\mathbf{A} = (\overline{z}, U_\varepsilon, V_\varepsilon)$, the full dynamics in the $2N$-dimensional space reduces to a three-dimensional (3D) space $(\beta^\star, \mu^\star, \nu^\star)$.

The general DynMaxEnt method is tested in the most challenging situation when the initially strong mutation suddenly changes to $4N\mu \ll 1$. The continuous DynMaxEnt method does not apply to this example for two reasons. First, when effective mutation drops below $4N\mu^\star = 1$ or $4N\nu^\star = 1$, the components of matrix $\mathbf{B}$ diverge because of singularities of the stationary allele frequency distribution. Second, even when the mutation rates are kept fixed at their terminal values ($\mu^\star = \mu$, $\nu^\star = \nu$), resulting in a reduced dynamics, the continuous DynMaxEnt method does not give correct dynamical predictions for small mutation rates. A closer look at the failures of the continuous DynMaxEnt method is provided in Appendix D.

The general method (with $\varepsilon = 1/2N$) gives a satisfactory estimate of the trait mean at the beginning of the adaptation process, as shown in Figure 1, while a quick drop in effective mutation, compensated by a change in the effective selection, results in a quick loss of polymorphism. When the equilibration of the effective mutation is too fast, the speed of the dynamics that are limited by effective mutation becomes slower than necessary and leads to underestimation of trait mean and $\langle U_\varepsilon \rangle$. This can be observed over the medium timescale $t/2N \sim 1$–10.

Over long timescales, the approximate dynamics accurately match the exact dynamics. When we employ an additional constraint on the mean heterozygosity, although no *a priori* selection is acting on it, we increase the number of degrees of freedom from three to four. While a general feature of variational approximations is that an increase in the number of constraints leads to an improvement in the performance of the approximation, this particular choice yields a match that is almost indistinguishable in all four observables (including the added $\langle pq \rangle$), as shown in Figure 1.

The dynamics of Lagrange multipliers in Figure 1 (E and F) demonstrates separation of timescales. An initial quick drop in mutation $\nu^\star$ and subsequently $\mu^\star$ to the vicinity of their steady states is compensated by the changes in $\beta^\star$ (and $h^\star$) and followed by a slow relaxation of $\beta^\star$ (and $h^\star$) to equilibrium once the mutation rates are barely changing. It is interesting that the dynamics of the effective forces differs significantly between cases E and F.

Example 2: Overdominance

Here we consider that the population dynamics is driven not only by directional selection but also by selection on heterozygosity, *i.e.*, overdominance. This yields stationary distributions that may have up to three modes: two modes represent peaks at the boundaries in the low-mutation regime, and an

intermediate peak represents an increased probability of heterozygous individuals. We consider an analogous example to Example 1 but add initially strong selection on heterozygosity (trimodal distribution of allele frequency) that switches to a purely directional selection.

Figure 2 shows an almost perfect agreement between the exact solution and the approximation despite the fact that the approximation is not built to capture rapid switches in the forces but, to the contrary, assumes that the forces change slowly. The effective forces do not capture the rapid change in the real forces on the system but show a slow–fast relaxation to their new steady states. Appendix H shows a complementary simulation when the initial (α_0) and terminal conditions (α_1) are switched. The system follows a different trajectory in the direction $\alpha_0 \rightarrow \alpha_1$ and $\alpha_1 \rightarrow \alpha_0$.

In all the simulations presented here, we use $2N\varepsilon = 1$; such a threshold ε corresponds to an allele frequency of a single individual out of a population of size $2N$. In Appendix F, we provide additional simulations to show that there exists an optimal threshold $N\varepsilon$ that minimizes the approximation error. Moreover, we show that in the worst-case scenario, where a strong change in selection is coincident with a low and changing mutation, the relative error of the approximation is on the order of 1%.

Example 3: Multiple loci with different effects

In this example, we show the applicability of the DynMaxEnt approximation for the evolution of quantitative traits that depend on multiple loci with different effects. The state space of the system of L independent loci contains essentially $(2N)^L$ degrees of freedom because each locus is characterized by an allele frequency distribution with $2N$ degrees of freedom. The reduction of the dimensionality to three (directional selection) or four (overdominance) degrees of freedom thus offers an immense simplification of the problem in which instead of tracking the full allele frequency distribution we are tracking just the dynamics of the Lagrange multipliers α^* corresponding to the underlying constraints.

How does the method perform for different distributions of effects? A simple case, in which all effects are the same ($\gamma_i = 1$), coincides with Example 2 because the dynamics become the same as for a single locus (Appendix I). We consider three distinct distributions of effects to demonstrate both the effect size distributions typically assumed and also extreme examples. We chose a uniform distribution in $[0, 2]$, exponential distribution with mean 1, and a bimodal distribution with many loci of small effect and a few loci of large effect. We chose all distributions to have $\langle \gamma \rangle = 1$ for an easier comparison.

Figure 3 shows a comparison between the exact dynamics and the general DynMaxEnt approximation. All forces, including forward and backward mutations, were initially perturbed to a state where $4N\mu = 4N\nu < 1$, the regime where the continuous approximation does not apply. When the distribution of effects was uniform or exponential, the approximation showed almost a perfect match with the true dynamics in all observed moments. Intuitively, one expects that when the distribution of effects is strongly bimodal, with many small effects and a few large effects, the approximation will perform badly. This is so because the large-effect loci dominate the quick response of the system, while the small-effect loci form a second wave of adaptation at a later time that is difficult to capture by the simple approximation schemes. Figure 3 shows these two
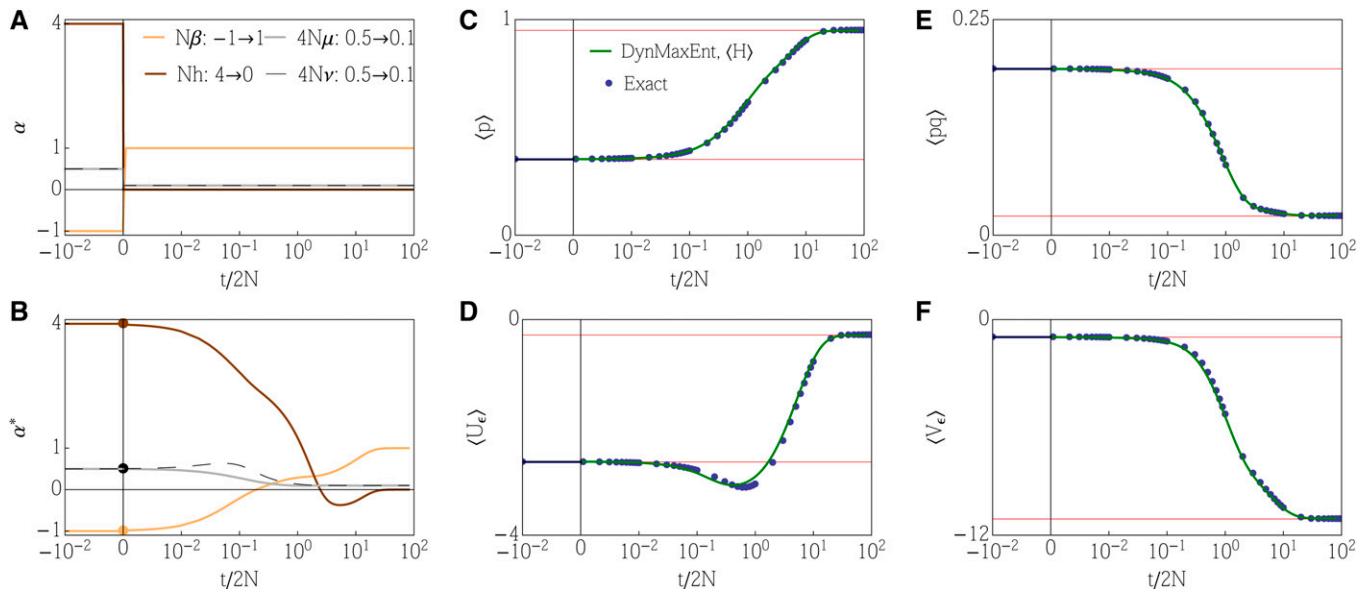
Figure 2 Example 2: details of the general DynMaxEnt method. (A) Changes in evolutionary forces, including a rapid decrease in mutation, complemented by a change in selection strength and a change in the selection strength on heterozygosity. (B) The effective forces α^*. (C–F) The response of the observed quantities $\langle p \rangle$, $\langle pq \rangle$, $\langle U_\varepsilon \rangle$, and $\langle V_\varepsilon \rangle$. Exact dynamics, obtained by numerically solving the diffusion equation using a discretization of space and time where the explicit transition matrix is known, are shown in blue dots, the general DynMaxEnt approximation is shown in green. We used $2N = 1000$ and $\varepsilon = 1/2N$.
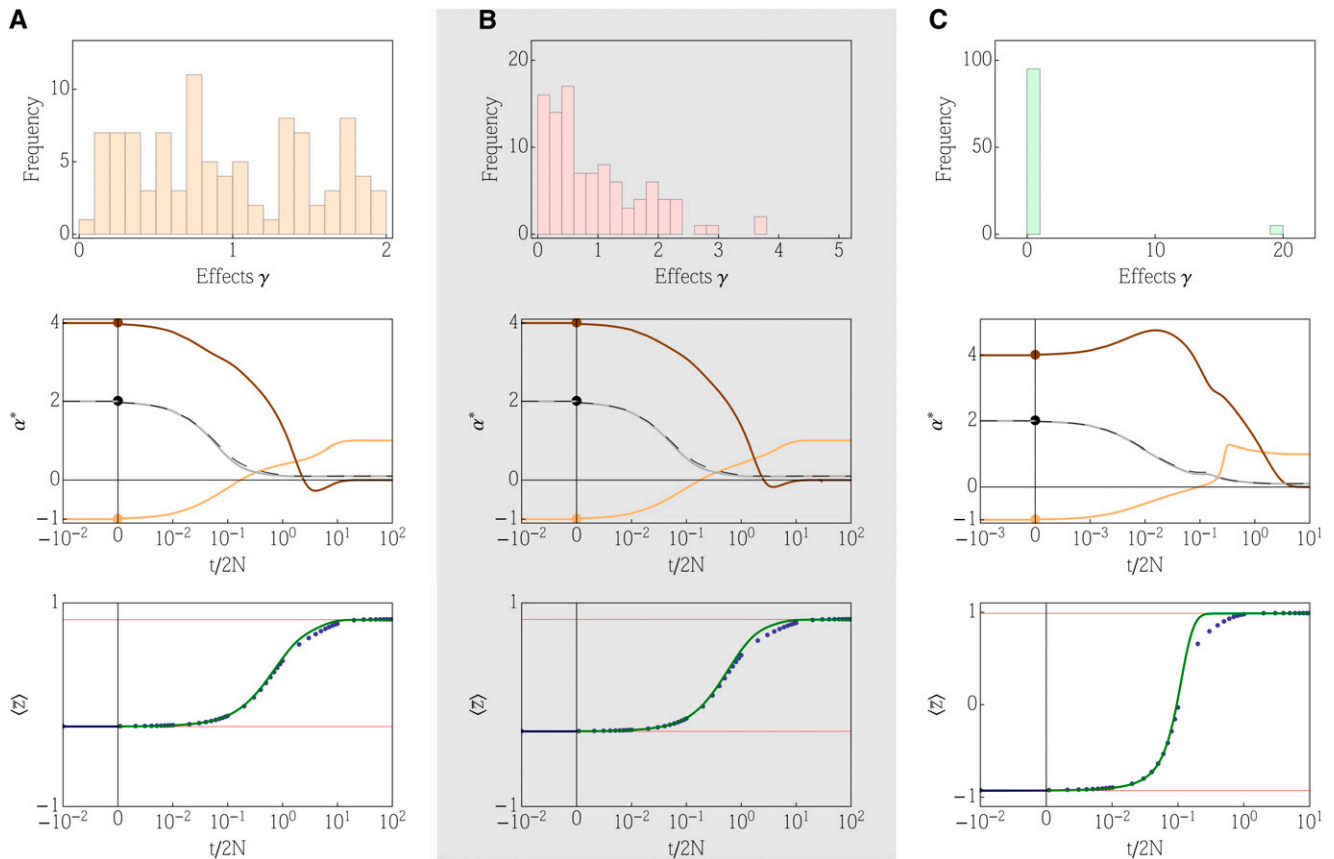
Figure 3 Example 3: dynamics of quantitative traits for 100 loci of different effects. The effects are randomly drawn from (A) uniform distribution in [0, 2], (B) exponential distribution with mean 1, and (C) deterministic effects where 95 of the loci have effects 0.01, while the remaining 5 loci have effects 19.81. The distributions were chosen to have a mean equal to 1 to be comparable with single-locus simulations. The rapid decrease in mutations is complemented by a change in selection strength and by a change in the selection strength on heterozygosity, as in Example 2. The true forces $\boldsymbol{\alpha}$ change at time $t = 0$ and draw the system out of equilibrium. The response of the trait mean $\langle \bar{z} \rangle$ is shown in blue dots, while the approximation is shown in green. The quality of the approximation for the remaining observables $\langle pq \rangle$, $\langle 2\log p \rangle$, and $\langle 2\log q \rangle$ can be found in Appendix J. Exact dynamics, obtained by numerically solving the diffusion equation using a discretization of space and time where the explicit transition matrix is known, are shown in blue dots. The general DynMaxEnt approximation is shown in green. We used $2N = 1000$ and $\varepsilon = 1/2N$.

timescales in the dynamics of Lagrange multipliers. The approximation of the observables still gives a solid match, but the second wave of adaptation due to small-scale effects is not perfectly accurate anymore.

Discussion

A central result in population genetics is that the stationary distribution of allele frequencies is the product of the neutral distribution and the mean fitness raised to the power of population size: $\psi \sim \phi \overline{W}^{2N}$ (Wright 1937; Kimura 1955b, 1964). This result can be interpreted via an optimization principle: selection constrains the expected values of the traits that determine fitness, but the allele frequency distribution is distorted as little as possible by this constraint—the distortion being measured by (minus) the relative entropy. Our approximation is to assume that this MaxEnt principle holds even away from equilibrium, with the approximate distribution of allele frequencies having the stationary functional *form* at all times. This provides a variational ansatz for

the diffusion equation and results in a set of dynamical equations for the parameters of the MaxEnt distribution.

This MaxEnt approximation can be justified in the limit of slowly changing conditions. Yet, provided that mutation rates are above a critical threshold ($4N\mu > 1$), it is remarkably accurate even in the worst case, when parameters change abruptly. Even for a single locus, this approximation gives a substantial reduction in complexity: the whole distribution is described by a small number of dynamical variables that correspond directly to the forces of mutation and selection. The method extends directly to multiple loci such that the joint distribution of allele frequencies can be approximated by a small number of variables.

This paper extends the approximation to low mutation rates ($4N\mu < 1$), where the original approximation breaks down completely. The problem is that when $4N\mu < 1$, populations are likely to be near fixation—*i.e.*, the distribution is concentrated near the borders. When mutation rate or population size changes, the distribution near the boundaries changes immediately, whereas the distribution in the interior changes much more slowly. To see why this is so, think of the probability that

a population carries one or a few copies of an allele. Because the lifetime of such rare alleles is short and is determined primarily by mutation and random reproduction, this probability changes rapidly with those processes and in the short term is independent of selection or of the total population size. In contrast, the distribution of polymorphic alleles in the interior changes slowly with changes in mutation rate because it takes a long time for new mutations to reach high frequency or for polymorphic alleles to be lost by drift.

The MaxEnt approximation for the expectations of a set of traits $\langle \mathbf{A} \rangle$ extends the "breeder's equation" to include random fluctuations, mutation, and an arbitrary relation between selected traits and the underlying genotype. It takes the form $\partial_t \langle \mathbf{A} \rangle = \mathbf{B}^*(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*)$ (13), where $\mathbf{B}^*$ is a generalization of the additive genetic covariance, $\boldsymbol{\alpha}$ are the actual forces of selection and mutation, and $\boldsymbol{\alpha}^*$ are the effective forces that would yield the same expectations $\langle \mathbf{A} \rangle$. Our extension is simply to modify $\mathbf{B}^*$ by truncating the distribution of allele frequencies within a distance $\varepsilon \sim 1/2N$ of the boundaries, thus suppressing its divergence for low mutation rates. The approximation is insensitive to the location of the truncation threshold ε. We give examples of directional selection and dominance; here we assume additivity across loci, but de Vladar and Barton (2011) show how the method applies to stabilizing selection, which induces pairwise epistasis between loci. We emphasize that the approximation can be applied without detailed knowledge of individual loci: only the distribution of allelic effects is needed. The approximation also can be generalized in principle to traits with linkage disequilibrium. The constraints for such a case also would involve pairwise measures, e.g., the correlations between loci.

Like the MaxEnt approximation, the infinitesimal model (Fisher 1919, p. 403) reduces the dimensionality of the dynamics by following trait values rather than individual alleles. However, these approaches are quite distinct. The infinitesimal model assumes that there are so many genes that the distribution of allele frequencies at each locus is hardly perturbed by selection; thus, the genetic variance that segregates within families remains approximately constant. This is equivalent to assuming that $N\beta$ on each allele is small (Robertson 1960). In contrast, the MaxEnt approximation can be applied to a single locus, with large $N\beta$, and predicts the change in genetic variance owing to selection. The infinitesimal model is broader, in that it describes the effects of linkage disequilibrium—though, because it assumes free recombination, these effects are only significant when selection on traits is strong; see Bulmer (1974) for an extension to allow linkage. However, if selection is weak enough that the population is at linkage equilibrium, the MaxEnt approximation will be more accurate than the infinitesimal model because it accounts for the effects of selection on the genetic variances. Of course, this does require knowledge of the distribution of effects of alleles and their interactions.

What are the possible applications of our results? First, our results allow us to tractably predict the temporal evolution of interesting macroscopic observables, such as the trait mean or heterozygosity, even when these are determined by an arbitrary function of a single locus or multiple loci under dynamically changing evolutionary forces. This is made possible by the drastic dimensionality reduction of the MaxEnt ansatz. The examples we presented here focused on exploring this "forward prediction" scenario in regimes where previously proposed approximation methods break down.

The MaxEnt approximation predicts the evolution of the allele frequency distribution—yet almost always we have only a single realization of the evolution of any one locus. However, whole-genome sequencing gives us information about the frequencies of alleles at very large numbers of loci. If these can be treated as independent realizations of a process common to all loci (or at least all loci in a functional class), then we can apply our method. Indeed, our assumptions are the same as those typically made in analyzing the distribution of frequencies of synonymous *vs.* nonsynonymous variants: each allele is taken to have an additive effect on fitness, drawn from some specific distribution. Usually, the distribution is assumed to be stationary. However, related species that have different effective population sizes (e.g., Loewe *et al.* 2006) or newly formed sex chromosomes (e.g., Zhou *et al.* 2013) require a time-dependent analysis of the kind proposed here.

Second, our results have consequences for inference of evolutionary forces from genomic and phenotypic data. The success of the MaxEnt approximation suggests that the allele frequency distribution remains close to the stationary form, even when selection, mutation, and population size are changing rapidly. This, in turn, suggests that it may be difficult to detect such changes from sequence data taken at a particular time point; note that the moments of the allele frequency distribution correspond directly to the distribution of genealogies. This is consistent with the finding that unless selection is very strong ($N\beta \gg 1$), it has only weak effects on genealogical structure (Williamson and Orive 2002; Barton and Etheridge 2004).

However, there are several reasons why this conclusion may be too pessimistic. First, even when the MaxEnt approximation accurately predicts the expectations of the observables, the underlying distribution may not necessarily be close to the stationary form [e.g., from Barton and de Vladar (2009, Figure 10)]. Second, when $4N\mu < 1$, sudden changes in population size and mutation rate cause immediate changes at the boundaries so that the distribution deviates from the stationary form—which is the problem that we address here. Indeed, it is relatively straightforward to detect strong population bottlenecks. Third, specific events, such as the sweep to fixation of a single mutation, can be detected. However, such events occur even when the ensemble is at a stationary state, so it still may be hard to find evidence for changing conditions. Tests that use an out-group or examine rates on a phylogeny (e.g., Kimura 1977; McDonald *et al.* 1991; Goldman and Yang 1994) can detect changes in selection but require multiple species and so are not covered by the arguments here. Moreover, our argument from MaxEnt only applies to freely recombining loci: additional information may come from patterns along the genome, which depend on linkage disequilibria and on rates of recombination. Lastly, if instead of data taken at a particular

time point we were provided with the temporal profile of changing observables $\langle \mathbf{A}(t) \rangle$, we could use our results to solve an inference problem and learn about the time courses of evolutionary forces $\boldsymbol{\alpha}$.

Third, our results, together with previous relevant theoretical work, allow us to interpret the evolutionary process in information-theoretic terms. What is the meaning of entropy, beyond being simply a tool for approximation? Minus the relative entropy is the Kullback-Leibler distance, a measure of divergence from the neutral distribution. If we include mutation in the base distribution ϕ (18), then minus the relative entropy measures the degree to which selection concentrates populations around states of high fitness. Following Kimura (1961), entropy changes can be seen as the information about its selective history that the population can transmit (Watkins 2008). Arguably, concentration around fit states is a better measure of adaptation than the increase in mean fitness: though fitness differences determine the rate of adaptation, they do not measure the outcome. Fitness may fluctuate, and absolute fitness must stay close to zero, on average, if populations are to persist. Mustonen and Lässig (2010) derive an intriguing relation between the gain in information (equal to the reduction in entropy) and the *fitness flux* Φ:

$$2N\mathbb{E}[\Phi] \geq -\Delta S \quad \text{where} \quad \Phi = \int_0^T \sum_i \beta_i \Delta p_i dt \qquad (28)$$

[(28) corrects a factor 2 error in Mustonen and Lässig (2010)]. This applies to a haploid population at linkage equilibrium, as assumed throughout this paper. Selection may change arbitrarily over time so that this relation gives a lower bound on the fitness flux Φ that is required to achieve a given gain in information (*i.e.*, reduction in entropy ΔS). If selection changes slowly—as required for our MaxEnt approximation to be accurate—then the inequality approaches an equality. The fitness flux can be separated into a component resulting from selection (which must be positive and equal to the additive variance in fitness V_W) and the remaining components, resulting from mutation and drift. Because forces other than selection are expected, on average, to act against adaptation, the latter component is negative, so the additive variance in fitness should set a bound on the rate of information gain (*i.e.*, $2NV_W \geq -\Delta S$).

Our approximation states that even out of equilibrium, the distribution of allele frequencies minimizes the information gain, subject to constraints on selected traits. By drastically reducing the dimensionality of the system to cover only the expectations of selected quantities, we can simplify expressions for the total fitness flux and variance in fitness over the evolutionary trajectory and therefore may be able to understand how these quantities limit the amount of information that can be accumulated by selection.

Acknowledgments

Literature Cited

Barton, N., and J. Coe, 2009 On the application of statistical physics to evolutionary biology. J. Theor. Biol. 259: 317–324.

Barton, N. H., and H. P. de Vladar, 2009 Statistical mechanics and the evolution of polygenic quantitative traits. Genetics 181: 997–1011.

Barton, N. H., and A. M. Etheridge, 2004 The effect of selection on genealogies. Genetics 166: 1115–1131.

Bialek, W., A. Cavagna, I. Giardina, T. Mora, E. Silvestri *et al.*, 2012 Statistical mechanics for natural flocks of birds. Proc. Natl. Acad. Sci. USA 109: 4786–4791.

Bulmer, M., 1974 Linkage disequilibrium and genetic variability. Genet. Res. 23: 281–289.

de Vladar, H. P., and N. H. Barton, 2011 The statistical mechanics of a polygenic character under stabilizing selection, mutation and drift. J. R. Soc. Interface 8: 720–739.

Ewens, W. J., 2012 *Mathematical Population Genetics 1: Theoretical Introduction* (Interdisciplinary Applied Mathematics, Vol. 27). Springer, New York.

Fisher, R. A., 1919 XV.—the correlation between relatives on the supposition of mendelian inheritance. Trans. R. Soc. Edinb. 52: 399–433.

Flegg, M. B., S. J. Chapman, and R. Erban, 2011 The two-regime method for optimizing stochastic reaction-diffusion simulations. *J. R. Soc. Interface* 9: 859–868.

Goeke, A., and S. Walcher, 2013 Quasi-steady state: Searching for and utilizing small parameters, pp. 153–178 in *Recent Trends in Dynamical Systems*. Springer, New York.

Goldman, N., and Z. Yang, 1994 A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol. Biol. Evol. 11: 725–736.

Hallatschek, O., 2011 The noisy edge of traveling waves. Proc. Natl. Acad. Sci. USA 108: 1783–1787.

Hick, P., and G. Stevens, 1987 Approximate solutions to the cosmic ray transport equation the maximum entropy method. Astron. Astrophys. 172: 350–358.

Hill, W. G., and M. Kirkpatrick, 2010 What animal breeding has taught us about evolution. Annu. Rev. Ecol. Evol. Syst. 41: 1–19.

Iwasa, Y., 1988 Free fitness that always increases in evolution. J. Theor. Biol. 135: 265–281.

Jaynes, E. T., 1957 Information theory and statistical mechanics. Phys. Rev. 106: 620.

Kimura, M., 1955a Solution of a process of random genetic drift with a continuous model. Proc. Natl. Acad. Sci. USA 41: 144.

Kimura, M., 1955b Stochastic processes and distribution of gene frequencies under natural selection. Cold Spring Harb. Symp. Quant. Biol. 20: 33–53.

Kimura, M., 1961 Natural selection as the process of accumulating genetic information in adaptive evolution. Genet. Res. 2: 127–140.

Kimura, M., 1962 On the probability of fixation of mutant genes in a population. Genetics 47: 713.

Kimura, M., 1964 Diffusion models in population genetics. J. Appl. Probab. 1: 177–232.

Kimura, M., 1977 Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. Nature 267: 275–276.

Loewe, L., B. Charlesworth, C. Bartolomé, and V. Nöel, 2006 Estimating selection on nonsynonymous mutations. Genetics 172: 1079–1092.

Lynch, M., B. Walsh *et al.*, 1997 *Genetics and Analysis of Quantitative Traits*, Vol. 1. Sinauer Associates, Sunderland, MA.

McDonald, J. H., M. Kreitman *et al.*, 1991 Adaptive protein evolution at the adh locus in Drosophila. Nature 351: 652–654.

Mora, T., A. M. Walczak, W. Bialek, and C. G. Callan, 2010 Maximum entropy models for antibody diversity. Proc. Natl. Acad. Sci. USA 107: 5405–5410.

Mustonen, V., and M. Lässig, 2007 Adaptations to fluctuating selection in Drosophila. Proc. Natl. Acad. Sci. USA 104: 2277–2282.

Mustonen, V., and M. Lässig, 2008 Molecular evolution under fitness fluctuations. Phys. Rev. Lett. 100: 108101.

Mustonen, V., and M. Lässig, 2010 Fitness flux and ubiquity of adaptive evolution. Proc. Natl. Acad. Sci. USA 107: 4248–4253.

Plastino, A., H. Miller, and A. Plastino, 1997 Minimum Kullback entropy approach to the Fokker–Planck equation. Phys. Rev. E 56: 3927–3934.

Pressé, S., K. Ghosh, J. Lee, and K. A. Dill, 2013 Principles of maximum entropy and maximum caliber in statistical physics. Rev. Mod. Phys. 85: 1115.

Prügel-Bennett, A., and J. Shapiro, 1997 An analysis of genetic algorithms using statistical mechanics. Physica D 104: 75–114.

Rattray, M., and J. L. Shapiro, 2001 Cumulant dynamics of a population under multiplicative selection, mutation, and drift. Theor. Popul. Biol. 60: 17–31.

Robertson, A., 1960 A theory of limits in artificial selection. Proc. R. Soc. Lond. B Biol. Sci. 153: 234–249.

Robinson, M., M. Flegg, and R. Erban, 2014 Adaptive two-regime method: application to front propagation. J. Chem. Phys. 140: 124109.

Schneidman, E., M. J. Berry, R. Segev, and W. Bialek, 2006 Weak pairwise correlations imply strongly correlated network states in a neural population. Nature 440: 1007–1012.

Segel, L. A., and M. Slemrod, 1989 The quasi-steady-state assumption: a case study in perturbation. SIAM Rev. 31: 446–477.

Sella, G., and A. E. Hirsh, 2005 The application of statistical physics to evolutionary biology. Proc. Natl. Acad. Sci. USA 102: 9541–9546.

Tkačik, G., O. Marre, D. Amodei, E. Schneidman, W. Bialek *et al.*, 2014 Searching for collective behavior in a large network of sensory neurons. PLoS Comut. Biol. 10: e1003408.

Tsimring, L. S., H. Levine, and D. A. Kessler, 1996 RNA virus evolution via a fitness-space model. Phys. Rev. Lett. 76: 4440.

Wakeley, J., 2008 *Coalescent Theory: An Introduction*, Vol. 1. Roberts & Company, Greenwood Village, CO.

Watkins, C., 2008 Selective breeding analysed as a communication channel: channel capacity as a fundamental limit on adaptive complexity, pp. 514–518 in *Proceedings of the 10th* International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC '08). IEEE Computer Society, Los Alamitos, CA.

Weigt, M., R. A. White, H. Szurmant, J. A. Hoch, and T. Hwa, 2009 Identification of direct residue contacts in protein-protein interaction by message passing. Proc. Natl. Acad. Sci. USA 106: 67–72.

Williamson, S., and M. E. Orive, 2002 The genealogy of a sequence subject to purifying selection at multiple sites. Mol. Biol. Evol. 19: 1376–1384.

Wright, S., 1931 Evolution in mendelian populations. Genetics 16: 97.

Wright, S., 1937 The distribution of gene frequencies in populations. Proc. Natl. Acad. Sci. USA 23: 307.

Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders *et al.*, 2010 Common snps explain a large proportion of the heritability for human height. Nat. Genet. 42: 565–569.

Zhou, Q., C. E. Ellison, V. B. Kaiser, A. A. Alekseyenko, A. A. Gorchakov *et al.*, 2013 The epigenome of evolving drosophila neo-sex chromosomes: dosage compensation and heterochromatin formation. PLoS Biol. 11: e1001711.

Communicating editor: J. Hermisson

Appendix A: Solution of MaxEnt by the Method of Lagrange Multipliers

The variational problem

$$\max_{\psi} S_{\mathrm{H}}[\psi] = - \int\limits_{[0,1]^{\mathrm{L}}} \psi(\mathbf{p},t)\log\left[\frac{\psi(\mathbf{p},t)}{\phi(\mathbf{p})}\right]d\mathbf{p} \tag{A1}$$

with $\phi(\mathbf{p}) = \prod_{i=1}^{\mathrm{L}}(p_i q_i)^{-1}$ subject to constraints

$$\int\limits_{[0,1]^{\mathrm{L}}} \psi(\mathbf{p})d\mathbf{p} = 1 \; , \qquad \int\limits_{[0,1]^{\mathrm{L}}} \mathbf{A}(p)\psi(\mathbf{p})d\mathbf{p} = \langle\mathbf{A}\rangle_{\mathrm{obs}*} \tag{A2}$$

[*e.g.*, for directional selection and asymmetrical mutation $\mathbf{A} = (\overline{z}, U, V)$ and $\boldsymbol{\alpha} = (\beta, \mu, \nu)$] can be solved by the method of Lagrange multipliers, yielding an unconstrained maximization of a Lagrangian

$$\mathcal{L}[\psi, \lambda, \boldsymbol{\alpha}] = S_{\mathrm{H}}[\psi] - \lambda\left[\int\limits_{[0,1]^{\mathrm{L}}} \psi(\mathbf{p})d\mathbf{p} - 1\right] - 2N\sum_{k}\alpha_k\left[\int\limits_{[0,1]^{\mathrm{L}}} A_k\psi(\mathbf{p})d\mathbf{p} - \langle A_k\rangle_{\mathrm{obs}}\right] \tag{A3}$$

with multipliers λ and $\boldsymbol{\alpha}$. The variational derivative of this function with respect to its argument ψ is

$$\delta\mathcal{L}[\psi, \lambda, \boldsymbol{\alpha}] = \int\limits_{[0,1]^{\mathrm{L}}} \delta\psi\left[\log\frac{\psi}{\phi} - 1 - \lambda - 2N\boldsymbol{\alpha}\cdot\mathbf{A}\right]d\mathbf{p} \tag{A4}$$

and leads to a solution

$$\overline{\psi}(\mathbf{p}) = \frac{\phi(p)}{\mathbb{Z}}e^{2N\boldsymbol{\alpha}\cdot\mathbf{A}} \tag{A5}$$

where the normalization $\mathbb{Z} = \exp(-1 - \lambda)$ and the Lagrange multipliers are such that observables are correctly matched. The distribution $\overline{\psi}(\mathbf{p})$ coincides with the stationary solution of the Fokker-Planck equation (1) provided that the $\boldsymbol{\alpha}$ values are the evolutionary forces (selection, mutation) and the $\mathbf{A}$ values are the traits associated with the underlying processes. The evolutionary forces appear in the constrained optimization as Lagrange multipliers $2N\boldsymbol{\alpha}$.

Appendix B: Matrices B and C in the Continuous DynMaxEnt Approximation

When $\boldsymbol{\alpha} = (\beta, h, \mu, \nu)$ and $\mathbf{A} = (\overline{z}, H, U, V)$, the explicit form of matrix $\mathbf{B}$ is

$$\mathbf{B}^* = \left\langle \sum_{k=1}^{\mathrm{L}}\begin{bmatrix} \gamma_k^2 p_k q_k & 2\gamma_k p_k^2 q_k(q_k - p_k) & \gamma_k q_k & -\gamma_k p_k \\ 2\gamma_k p_k^2 q_k(q_k - p_k) & 2p_k^3 q_k(q_k - p_k)^2 & 2p_k q_k(q_k - p_k) & -2p_k^2(q_k - p_k) \\ \gamma_k q_k & 2p_k q_k(q_k - p_k) & 2\frac{q_k}{p_k} & -2 \\ -\gamma_k p_k & -2p_k^2(q_k - p_k) & -2 & 2\frac{p_k}{q_k} \end{bmatrix}\right\rangle \tag{B1}$$

as implemented in the Mathematica notebook (File S1). Because the expectation $\langle f(\mathbf{p})\rangle = \int_{[0,1]^{\mathrm{L}}}f(\mathbf{p})\overline{\psi}(\mathbf{p})d\mathbf{p}$ is taken over the stationary distribution $\overline{\psi}(\mathbf{p})$, which depends on the effective parameters $\boldsymbol{\alpha}^*$, the matrix $\mathbf{B}^*$ is also a function of the effective evolutionary forces. The components of matrix $\mathbf{B}^*$ may be calculated using special functions, as in Barton and de Vladar (2009).

Similarly, the components of the matrix $\mathbf{C}^*$ can be written as

$$C_{i,j}^* = \mathrm{Cov}(A_i, A_j) = \langle A_i A_j\rangle - \langle A_i\rangle\langle A_j\rangle$$

where $A_i(\mathbf{p})$ are functions of the microscopic allele frequencies, and after averaging over the stationary distribution $\overline{\psi}(\mathbf{p})$, only the dependence on the effective forces $\boldsymbol{\alpha}^*$ remains. Thus, the right-hand side of vector equation (14) is solely a function of effective evolutionary forces $\boldsymbol{\alpha}^*$, forming a system of ordinary differential equations of dimension equal to the number of observables.

Appendix C: Discrete Dynamics in the Limit of Small $N\mu$

As $4N\mu$ and $4N\nu$ become very small, the probability distribution (7) becomes concentrated at the boundaries. The populations then switch between fixation for the favorable and deleterious alleles at a given locus and can be described by the fraction P, $Q = 1 - P$, of populations fixed (or nearly fixed) for each allele:

$$\frac{dP}{dt} = \lambda_+ Q - \lambda_- P \tag{C1}$$

The probability of fixation for the favorable allele is $\hat{P} = [\nu/\mu + \exp(-4N\beta)]^{-1}$ (Kimura 1962), and the rates of substitution of alleles by their counterparts, λ_+ and λ_-, are

$$\lambda_- = 4N\nu\beta \frac{e^{-4N\beta}}{1 - e^{-4N\beta}} \quad \text{and} \quad \lambda_+ = 4N\mu\beta \frac{1}{1 - e^{-4N\beta}} \tag{C2}$$

Hence, the exact dynamics in the regime of small mutations have a form

$$4N\frac{dP}{dt} = 4N\beta \frac{4N\mu Q - 4N\nu e^{-4N\beta} P}{1 - e^{-4N\beta}} \tag{C3}$$

How does the standard MaxEnt approximation compare in the limit of low mutation rates? We keep mutation rates fixed and follow a single variable $P = \langle p \rangle$; we define the complementary variable β^* as the selection that gives P at stationarity. In the limit of low mutation rates,

$$p = \frac{\int_0^1 p \cdot p^{4N\mu-1} q^{4N\nu-1} e^{4N\beta^*(2p-1)} dp}{\int_0^1 p^{4N\mu-1} q^{4N\nu-1} e^{4N\beta^*(2p-1)} dp} \approx \frac{\dfrac{1}{4N\nu}}{e^{-4N\beta^*}\dfrac{1}{4N\mu} + \dfrac{1}{4N\nu}} \tag{C4}$$

implying

$$e^{4N\beta^*} = \frac{4N\nu}{4N\mu} \frac{P}{Q} \tag{C5}$$

The same equilibrium formula is given by the ratio of substitution rates $P/Q = \lambda_+/\lambda_-$. Under the MaxEnt approximation, the rate of change is

$$4N\frac{dP}{dt} = \langle pq \rangle (4N\beta - 4N\beta^*) \tag{C6}$$

Moreover, in the limit of small mutations,

$$\langle pq \rangle = \frac{\int_0^1 p^{4N\mu} q^{4N\nu} e^{4N\beta^*(2p-1)} dp}{\int_0^1 p^{4N\mu-1} q^{4N\nu-1} e^{4N\beta^*(2p-1)} dp} \approx \frac{e^{4N\beta^*} - 1}{4N\beta^*\left(\dfrac{1}{4N\mu} + \dfrac{e^{4N\beta^*}}{4N\nu}\right)} \tag{C7}$$

where this expression has been obtained by computing contributions to the integrals separately for each of the boundaries $[0, \delta]$ and $[1 - \delta, 1]$, for $\delta \ll 1$ and for $4N\mu, 4N\nu \ll \delta$. Note that integration by parts has been used, resulting in the presence of the term $4N\beta^*$ in the denominator of (C7). Equation (C6) then becomes

$$4N\frac{dP}{dt} = \frac{4N\nu P - 4N\mu Q}{4N\beta^*}(4N\beta - 4N\beta^*) = \left[\frac{4N\beta}{\log\left(\dfrac{4N\nu}{4N\mu}\dfrac{P}{Q}\right)} - 1\right](4N\nu P - 4N\mu Q) \tag{C8}$$

Thus, the MaxEnt approximation to the full distribution does not converge to the exact dynamics (C3) as mutation rates become small. Nevertheless, the dynamics are approximated quite well, provided that selection is not too strong (*e.g.*, $|4N\beta| < 2$). The MaxEnt approximation greatly underestimates the rate of change at the margins and gives no effect of selection at the extreme allele frequencies (see Figure C1, left).
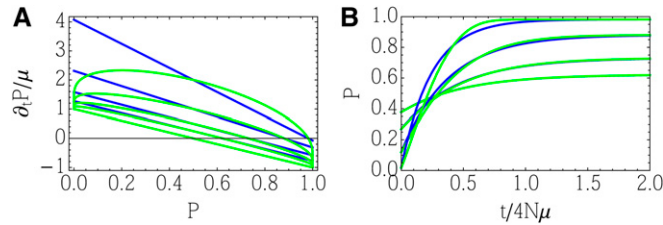
Figure C1 Discrete MaxEnt approximation. (A) The exact rate of change, obtained from the discrete model (C3) (blue), compared with the MaxEnt approximation (C8) (green) in the limit of small $4N\mu = 4N\nu$; $N\beta = 0, 0.5, 1, 2, 4$ (bottom to top) and $2N = 1000$. Rates are scaled relative to the neutral mutation rate, assumed to be symmetrical ($\mu = \nu$). (B) Dynamics following an abrupt switch from $4N\beta = -0.5, -1, -2, -4$ to the reverse.

However, the equilibria necessarily agree: the exact and the approximate rates of change are zero at the same point. The right plot in Figure C1 shows the exact $P(t)$ and the MaxEnt approximation; these are close to $4N\beta = \pm 2$, but the solution becomes poor for $4N\beta = 4$ (upper pair of curves). The MaxEnt solution is accurate for $4N\beta \leq 2$ because P remains within the interior $[1/(1 + e^2) = 0.112 < P < 0.888)$, where the rate of change is approximated well.

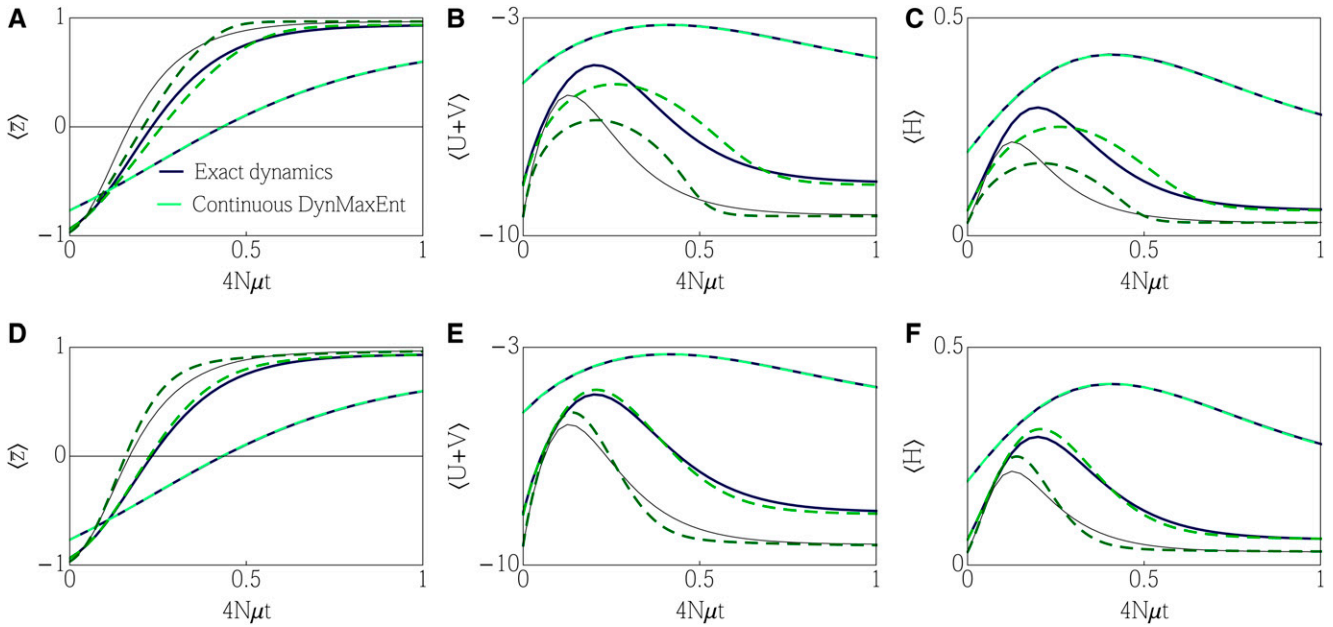## Appendix D: Failure of the Continuous DynMaxEnt for $4N\mu, 4N\nu < 1$

The log mean fitness typically will be a sum over moments of allele frequencies. For example, a selection gradient β on a trait with mean $\bar{z} = 2\sum_i \gamma_i(p_i - q_i)$ will introduce a component where $\alpha_W = \beta$ and $A_W = \bar{z}$; a model with dominance requires $\alpha_W = (\beta, h)$ and $A_W = (\bar{z}, H)$, and epistasis introduces mixed second-order moments of the allele frequencies. Thus, the matrix **B** is an expectation over polynomial functions of allele frequencies and is well behaved.

In contrast, the elements of **B** that describe mutation diverge when $4N\mu < 1$ or $4N\nu < 1$. To see this, consider a single locus for which $A = (\bar{z}, H, U, V) = (2p - 1, 2pq, 2\log p, 2\log q)$, and the elements $B_{3,3}$, $B_{3,4}$, and $B_{4,4}$ are $\langle 2q/p \rangle, \langle -2 \rangle, \langle 2p/q \rangle$ (B1). Thus, $B_{3,3}$ diverges when $4N\mu < 1$, and $B_{4,4}$ diverges when $4N\nu < 1$. If we can assume that mutation rates are fixed, then we can avoid the difficulty either by fixing the mutation rates always at their actual values (*i.e.*, $\mu = \mu^*$, $\nu = \nu^*$) or by choosing a reference distribution that includes mutation, *e.g.*, $\phi = \prod_k p_k^{4N\mu-1} q_k^{4N\nu-1}$, and dropping the observables U and V. These two approaches are equivalent because fixing the mutation rate leads to $\partial_t \langle \bar{z} \rangle = B_{1,1}(\beta - \beta^*) + B_{1,2}(h - h^*)$, leading to the same dynamics for β^* and h^* as if the reference distribution included mutation.

We first explore the continuous DynMaxEnt approximation for $4N\mu > 1$ and compare its accuracy with that for $4N\mu < 1$. We study the worst-case scenario when a selection suddenly changes sharply from $N\beta = -4$ to $N\beta = 4$; this is by no means an adiabatic change. Mutations equal $4N\mu = 4N\nu = 2, 1/2, 1/4$, where the first choice ($4N\mu > 1$) allows a full continuous DynMaxEnt approximation including constraints on $\langle U \rangle$ and $\langle V \rangle$, while the second and third choices require a fixed instantaneous mutation rate and constraints on $\langle U \rangle$ and $\langle V \rangle$ dropped to keep the entries of the additive covariance matrix **B** finite. Figure D1 (top row) shows the predicted observables $\langle \bar{z} \rangle$, $\langle U + V \rangle$, and $\langle H \rangle$ estimated by the continuous DynMaxEnt method for each of the mutation rate alternatives compared with the exact solution, while keeping the heterozygosity fixed throughout the simulation (not employing the constraint on $\langle 2pq \rangle$). The method is accurate for $4N\mu = 2$ (where the mutation rate also changes dynamically) but shows significant deviations from the true dynamics for $4N\mu = 4N\nu = 1/2, 1/4$.

We also show simulations (Figure D1, D–F) in which the selection on heterozygosity h is treated as a dynamical force, employing an additional constraint on $\langle H \rangle$. This increases the number of degrees of freedom in the approximate dynamics and affects the accuracy of the approximation. For instance, the decrease in performance for $4N\mu$, $4N\nu < 1$ is partially caused by losing two degrees of freedom by fixing μ and ν. However, an increase in accuracy in Figure D1 (bottom row) is caused by
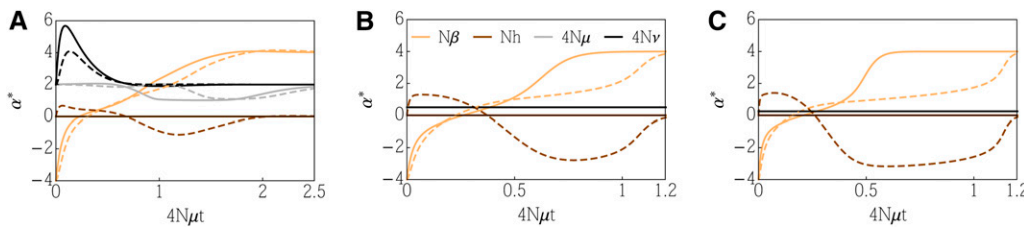
Figure D1 Failure of the continuous DynMaxEnt approximation for strong selection. Dynamical response to a fast change in selection $N\beta = -4$ to $N\beta = 4$ for symmetrical mutation rates that are unperturbed and $2N = 1000$. Each panel shows three simulations with mutation rates $4N\mu = 4N\nu \in \{2, 1/2, 1/4\}$ (light-green to dark-green) where the exact observables (black) are compared with the continuous DynMaxEnt approximation (green, dashed). (A–C) Simulation with constraints $\langle \bar{z} \rangle$, $\langle U \rangle$, and $\langle V \rangle$ when $4N\mu, 4N\nu > 1$ and a single constraint $\langle \bar{z} \rangle$ when $4N\mu, 4N\nu < 1$. (D–F) Simulation with constraints $\bar{z}$, $\langle H \rangle$, $\langle U \rangle$, and $\langle V \rangle$ when $4N\mu, 4N\nu > 1$ and two constraints $\langle \bar{z} \rangle$ and $\langle H \rangle$ when $4N\mu, 4N\nu < 1$.

introducing additional constraints on $\langle 2pq \rangle$. This is visible both for $4N\mu > 1$, where the number of Lagrange multipliers was increased from three to four, and for $4N\mu < 1$, where the constraint on $\langle \bar{z} \rangle$ was complemented by the second constraint on a mean heterozygosity $\langle H \rangle = \langle 2pq \rangle$.

Figure D2 presents the dynamics of Lagrange multipliers in the example in Figure D1, where a sudden change in selection from $N\beta = -4$ to $N\beta = 4$ was applied to three systems with different mutation rates $N\mu, N\nu \in \{2, 1/2, 1/4\}$. While in the superthreshold regime $4N\mu, 4N\nu > 1$, the mutation is allowed to change in the continuous DynMaxEnt approximation; in the subthreshold regime $4N\mu, 4N\nu < 1$, it is fixed.

Figure D2 suggests that increasing the number of constraints; which increases the dimensionality of the problem, is associated with a slower convergence to a steady state and a separation of timescales. This is visible both for $4N\mu > 1$ (left), where the number of Lagrange multipliers was increased from three to four, and for $4N\mu < 1$, where the constraint on $\langle \bar{z} \rangle$ was complemented by the second constraint on a mean heterozygosity $\langle H \rangle = \langle 2pq \rangle$.

Next, suppose that the scaled mutation rate is initially high enough that the stationary distribution is concentrated in the interior but then abruptly falls below the threshold at which populations are typically near fixation: $4N\mu = 4N\nu$ falls from 2 to 0.1. (Note that a fall in $N\mu$ also could be due to a fall in population size rather than in mutation rate.) We also assume that selection changes abruptly at the same time (*i.e.*, $N\beta$ changes from -1 to 1) in order to give the most challenging example: errors in estimating $\langle pq \rangle$ will be reflected in errors in the rate of change of the mean. Immediately after the fall in scaled mutation rate, probability accumulates at the boundary and develops a boundary layer of a form $\sim p^{4N\mu-1}q^{4N\nu-1}$ that agrees

Figure D2 Effective coefficients α^* of the continuous DynMaxEnt approximation for the scenario of Figure D1. (A) $4N\mu = 4N\nu = 2$, (B) $4N\mu = 4N\nu = 1/2$, and (C) $4N\mu = 4N\nu = 1/4$ are compared between the simulation without (solid line) *vs.* with (dashed line) the heterozygosity as a degree of freedom. The population size is $2N = 1000$.

with the stationary form, but polymorphism decays more slowly, so the full distribution is not from the stationary class. The continuous DynMaxEnt method fails to capture the true dynamics because the mutation rates have to be instantly adjusted to the terminal values.

Figure D3 shows that three measures of diversity ($\langle U \rangle, \langle V \rangle, \langle pq \rangle$) change rapidly, with $\langle U \rangle$ and $\langle V \rangle$ falling most rapidly because they are more sensitive to the rapid changes near the boundary than $\langle pq \rangle$. Note that $\langle U \rangle = \langle 2 \log p \rangle$ falls until $t \sim 1/2N$ because probability accumulates close to $p = 0$ (note the log scale in Figure D3) but then increases again more slowly as favorable mutations substitute, transferring probability away from $p \sim 0$. The mean changes rapidly and substantially while
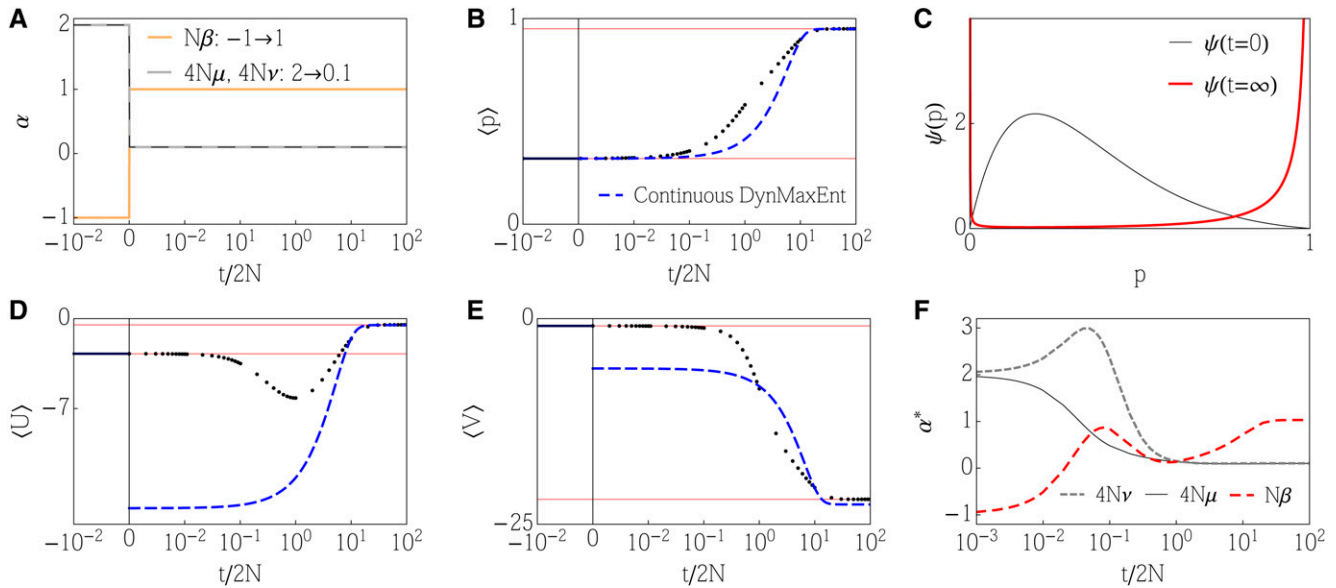
Figure D3 Failure of continuous DynMaxEnt approximation for changing mutations $4N\mu < 1$ and $2N = 1000$. (A) Changes in evolutionary forces that draw the system out of equilibrium include a rapid decrease in mutation, complemented by a change in selection strength. (B, D, and E) The response of the observed quantities $\langle p \rangle$, $\langle U \rangle$, and $\langle V \rangle$. (C) The initial and final stationary allele frequency distributions. (F) The effective forces $\boldsymbol{\alpha}^*$.

heterozygosity is still high but then changes more slowly after $t \sim 1/2N$, when the genetic variance is low and selection is limited by the influx of new mutations.

The changes in effective parameters $\boldsymbol{\alpha}^*$ are shown in Figure D3F. Because we assume that the population is initially in its stationary state, these parameters necessarily begin at their actual values ($4N\mu^*, 4N\nu^* = 2, N\beta^* = -1/2$). After the mutation rate decreases, the probability of being fixed increases rapidly, and therefore, $4N\mu^*$ and $4N\nu^*$ fall quickly, approaching their new values by $t/2N \sim 0.1$; μ^* falls first because the probability of being near $p = 0$ increases faster than the probability of being near $p = 1$. Over this time, $N\beta^*$ changes even while the mean $\langle p \rangle$ hardly changes in order to compensate for changing $N\mu$ and $N\nu$. At later times, when the effective mutation rates are constant and close to their actual (low) rates, $N\beta^*$ increases as the mean changes.

Fixing the effective mutation rates makes DynMaxEnt unable to capture transient properties of the adaptation process. The dynamics of the mean allele frequency follow $p' = 2\langle pq \rangle(\beta - \beta^*)$, where β^* can be obtained at each time as a value that gives the current mean allele frequency given $\mu^* = \mu$ and $\nu^* = \nu$. The effective directional selection can be used further to compute the change of observables, as displayed in Figure D3 (dashed blue). This continuous DynMaxEnt approximation fails to capture transient dynamics of the observed quantities but still converges to the correct state. If the exact dynamics are followed for a short time from the switch of the evolutionary forces and only then is the continuous DynMaxEnt approximation initialized, the approximation would more closely agree with the actual observables. This is so because the real dynamics of the trait mean slow down in time as the system gets close to an equilibrium and therefore are better captured by an approximate process whose speed of adaptation is limited by the small fixed mutation rate.

Appendix E: Derivation of the General DynMaxEnt Approximation for Low Mutation from the Diffusion Equation 1

We consider a directional selection acting on a single locus. The extension to multiple loci with different effects and overdominance is straightforward and discussed later.

The failure of the continuous DynMaxEnt method, described in Appendices C and D, requires a special treatment of the boundaries of the allele frequency domain to avoid divergence of the method owing to singularities in matrix **B**. A naive resolution is to completely disregard the boundaries and simply to truncate the allele frequency distribution. This *ad hoc* method, despite resolving problems with divergence of matrix **B**, would not capture the dynamics at the boundaries very accurately, particularly in the small-mutation regime. Thus, we approach the problem in a more pedantic way.

First, we split the domain to the interior and the boundaries and propose a dynamical model that captures all essential processes. The new system depends on a truncation parameter ε representing the width of the boundary layer. The stationary distribution of the system still can be represented as a solution of the variational problem with two additional constraints and Lagrange multipliers related to the boundary masses. The DynMaxEnt method can be derived for this expanded system. However, the resulting macroscopic dynamics of the effective forces involve inclusion of the coupling terms at $p = \varepsilon, 1 - \varepsilon$, which are microscopic quantities; these are not in general accessible. Therefore, we introduce an approximation that does not directly prescribe the probability fluxes but circumvents this by setting the total mass at the boundaries consistently with the continuous stationary distribution. This step is crucial for our approximation. The effect is that the expanded dynamics are brought back to a dimension where only the mutation and selection forces are followed.

The derivation of the general DynMaxEnt approximation for a system where the stationary solution is a discontinuous function

$$\psi_\varepsilon(p,t) = \begin{cases} \psi_\varepsilon^1(p,t) & p < \varepsilon \\ \psi_\varepsilon^2(p,t) & \text{else} \\ \psi_\varepsilon^3(p,t) & p > 1 - \varepsilon \end{cases} \tag{E1}$$

consists of the following steps:

1. Set up a piecewise-defined diffusion dynamics with the appropriate stationary solution.
2. Use the stationary solution as an ansatz for a solution of the diffusion problem, allowing only the Lagrange multipliers to change in time.
3. Derive dynamics of the means of observables from this system and reduce it to a closed dynamical system for Lagrange multipliers.

1. Diffusion equation (1) can be written in the flux form

$$\frac{\partial \psi}{\partial t} = -\frac{\partial}{\partial p} J[\psi, p] \tag{E2}$$

where the flux J is defined as

$$J[\psi, p] = [\mathcal{M}_i(p_i)\psi] - \frac{1}{2}\frac{\partial}{\partial p}[\mathcal{V}_i(p_i)\psi] \tag{E3}$$

We set up the dynamical system for $\psi_\varepsilon(p,t)$ separately in the three regions of the domain:

$$\frac{\partial \psi_\varepsilon}{\partial t} = -\frac{\partial}{\partial p} J_\varepsilon[\psi_\varepsilon, p] = \begin{cases} -\dfrac{\partial}{\partial p} J_\varepsilon^1[\psi_\varepsilon^1, p] & p \in [0, \varepsilon) \\[2mm] -\dfrac{\partial}{\partial p} J_\varepsilon^2[\psi_\varepsilon^2, p] & p \in (\varepsilon, 1 - \varepsilon) \\[2mm] -\dfrac{\partial}{\partial p} J_\varepsilon^3[\psi_\varepsilon^3, p] & p \in (1 - \varepsilon, 1] \end{cases} \tag{E4}$$

where J_ε are the fluxes in the system with split domain (E5)

$$J_\varepsilon^1[\psi^1, p] = [(\beta \varepsilon q - v\varepsilon)\psi_\varepsilon^1] - \frac{1}{4N}\frac{\partial}{\partial p}[\varepsilon q \psi_\varepsilon^1]$$

$$J_\varepsilon^2[\psi^2, p] = [(\beta pq - vp + \mu q)\psi_\varepsilon^2] - \frac{1}{4N}\frac{\partial}{\partial p}[pq \psi_\varepsilon^2] \tag{E5}$$

$$J_\varepsilon^3[\psi^3, p] = [(\beta p\varepsilon + \mu\varepsilon)\psi_\varepsilon^3] - \frac{1}{4N}\frac{\partial}{\partial p}[p\varepsilon \psi_\varepsilon^3]$$

and where the coupling between regions, which influences the probability of extreme allele frequencies, is set as

$$J_\varepsilon^1\left[\psi_\varepsilon^1, 0\right] = 0 \tag{E6}$$

$$J_\varepsilon^3\left[\psi_\varepsilon^3, 1\right] = 0 \tag{E7}$$

$$J_\varepsilon^1\left[\psi_\varepsilon^1, \varepsilon\right] = J_\varepsilon^2\left[\psi_\varepsilon^2, \varepsilon\right] = J[\psi, \varepsilon] \tag{E8}$$

$$J_\varepsilon^2\left[\psi_\varepsilon^2, 1-\varepsilon\right] = J_\varepsilon^3\left[\psi_\varepsilon^3, 1-\varepsilon\right] = J[\psi, 1-\varepsilon] \tag{E9}$$

While the conditions (E6 and E7) at $p = 0, 1$ ensure that the probability mass does not dissipate through the boundaries, the conditions at $p = \varepsilon, 1 - \varepsilon$ set the fluxes in accordance with the fluxes of the original continuous dynamics (1), which are not explicitly known. This results in the same stationary solution in the bulk as in the original diffusion process. Moreover, the probability masses at the boundaries also will agree with the standard problem because the fluxes are identical. The stationary solution, computed for $p \in [0, \varepsilon]$ as a solution of the problem $J_\varepsilon^1[\psi^1, p] = 0$ and analogously for subdomains $(\varepsilon, 1 - \varepsilon)$ and $(1 - \varepsilon, 1]$, has a form (17). Moreover, setting the total mass at each subdomain equal to the corresponding mass of the stationary solution (7) to leading order implies the choice of constant prefactors in (22).

2. The stationary distribution $\overline{\psi}_\varepsilon(p|\boldsymbol{\alpha})$ in (17) depends on five parameters (for directional selection), including κ and ρ related to boundary masses. These parameters relate to a set of observables $\mathbf{A}^{\text{full}} = (\overline{z}, U_\varepsilon, V_\varepsilon, \chi_{[0,\varepsilon]}, \chi_{[1-\varepsilon,1]})$ as the corresponding Lagrange multipliers $2N\boldsymbol{\alpha}^{\text{full}} = 2N(\beta, \mu, \nu, \kappa, \rho)$. We derive the full DynMaxEnt method for this extended problem. Only after the derivation of the full dynamics will we incorporate the relationship between κ, μ and ρ, ν. We use an ansatz

$$\psi_\varepsilon(p, t) = \overline{\psi}_\varepsilon(p|\boldsymbol{\alpha}^\star) + \delta(p, t) \tag{E10}$$

for the solution of (E4) that expresses the nonstationary solution of the equation as a sum of a stationary distribution, with effective Lagrange multipliers $\boldsymbol{\alpha}^\star$ changing in time and δ representing the deviation from the stationary form. First, we express the left-hand side of (E4) using a chain rule:

$$\frac{\partial \overline{\psi}_\varepsilon}{\partial t} = \frac{\partial \overline{\psi}_\varepsilon}{\partial \beta^\star}\frac{\partial \beta^\star}{\partial t} + \cdots + \frac{\partial \overline{\psi}_\varepsilon}{d\rho^\star}\frac{\partial \rho^\star}{\partial t} + \frac{\partial \delta}{\partial t}$$

Next, we calculate each summand and, for clarity, drop the notation in the exponent of $\boldsymbol{\alpha}^{\text{full}}$:

$$\frac{\partial \overline{\psi}_\varepsilon}{\partial \alpha_k^\star} = -\frac{1}{\mathbb{Z}}\frac{\partial \mathbb{Z}}{\partial \alpha_k^\star}\overline{\psi}_\varepsilon + 2NA_k\overline{\psi}_\varepsilon = 2N[A_k - \langle A_k\rangle]\overline{\psi}_\varepsilon$$

where $\langle A_k\rangle$ arises from differentiating $\mathbb{Z}$ and A_k from differentiating $e^{2N\boldsymbol{\alpha}\cdot\mathbf{A}}$ with respect to $\boldsymbol{\alpha}_k^\star$. This leads to

$$\frac{\partial \overline{\psi}_\varepsilon}{\partial t} = 2N\overline{\psi}_\varepsilon\sum_{k=1}^5[A_k - \langle A_k\rangle]\frac{\partial \alpha_k^\star}{\partial t} + \frac{\partial \delta}{\partial t}$$

Next, we express the right-hand side of (E4) by artificially adding an expression $J_\varepsilon[\overline{\psi}_\varepsilon, p|\boldsymbol{\alpha}^\star]$. This term represents dynamics where the forces, applied to the system, equal the effective forces. Given stationarity, the term equals zero. The reason for its inclusion is a cancellation of the diffusive part of the equation that does not depend on $\boldsymbol{\alpha}$. This leads to a simple outcome where each term contains the difference between the true and the effective force:

$$-\frac{\partial}{\partial p}J_\varepsilon\left[\overline{\psi}_\varepsilon, p|\boldsymbol{\alpha}\right] = -\frac{\partial}{\partial p}J_\varepsilon\left[\overline{\psi}_\varepsilon, p|\boldsymbol{\alpha}\right] + \frac{\partial}{\partial p}J_\varepsilon\left[\overline{\psi}_\varepsilon, p|\boldsymbol{\alpha}^\star\right]$$

$$= -\frac{\partial}{\partial p}\left[\frac{\sigma_\varepsilon(p)}{2}\sum_{k=1}^5\left(\alpha_k - \alpha_k^\star\right)\frac{\partial A_k}{\partial p}\overline{\psi}_\varepsilon\right]$$

Equation (E4) thus becomes

$$2N\overline{\psi}_\varepsilon \sum_{k=1}^{5} [A_k - \langle A_k \rangle] \frac{\partial \alpha_k^*}{\partial t} + \frac{\partial \delta}{\partial t} = -\frac{\partial}{\partial p}\left[\frac{\sigma_\varepsilon(p)}{2} \sum_{k=1}^{5}\left(\alpha_k - \alpha_k^*\right) \frac{\partial A_k}{\partial p} \overline{\psi}_\varepsilon\right] - \frac{\partial J_\varepsilon[\delta, p]}{\partial p} \tag{E11}$$

where the function $\sigma_\varepsilon(p) = pq$ in the bulk and εq and $p\varepsilon$ when $p < \varepsilon$ and $p > 1 - \varepsilon$, respectively.

3. We then multiply the equation by A_j and average via allele frequency distribution. The goal is to find effective forces $\boldsymbol{\alpha}^*$ such that the error terms vanish; *i.e.*, the projection of the full dynamics to the space of macroscopic quantities $\mathbf{A}$ is closed. Because there are k constraints and the same number of forces, such $\boldsymbol{\alpha}$ values in principle exist and are unique. The most crucial implication is that the approximation of macroscopic quantities is forced to coincide with the exact values $\mathbf{A}$ while the quasi-stationarity is valid. The equation becomes

$$2N \sum_{k=1}^{5} \frac{\partial \alpha_k^*}{\partial t} \int_0^1 [A_k A_j - \langle A_k \rangle A_j] \overline{\psi}_\varepsilon dp = -\sum_{k=1}^{5} \int_0^1 A_j \frac{\partial}{\partial p}\left[\frac{\sigma_\varepsilon(p)}{2}\left(\alpha_k - \alpha_k^*\right) \frac{\partial A_k}{\partial p} \overline{\psi}_\varepsilon\right] dp \tag{E12}$$

Next, we use integration by parts on the right-hand side of the equation. This introduces boundary terms, *i.e.*, terms evaluated at $p = 0, \varepsilon, 1 - \varepsilon, 1$ coming from the integration by parts. We neglected these terms by assuming a rapid convergence to the stationary distribution and instantaneous adjustment of the fluxes at the boundaries of the subdomains to their stationary values. This stationarity assumption is adopted to avoid dependence of the approximation on the microscopic details of the distribution. This leads to a relationship between the moment dynamics and dynamics of Lagrange multipliers in the full model:

$$2N\mathbf{C}_\varepsilon^{\text{full}} \frac{\partial \boldsymbol{\alpha}^*}{\partial t} = \mathbf{B}_\varepsilon^{\text{full}}(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*) \tag{E13}$$

where $\boldsymbol{\alpha} = \boldsymbol{\alpha}^{\text{full}}$ and symmetrical matrices

$$\mathbf{B}_\varepsilon^{\text{full}} = \left\langle \frac{\sigma_\varepsilon(p)}{2} \frac{\partial A_j}{\partial p} \frac{\partial A_k}{\partial p} \right\rangle \quad \text{and} \quad \mathbf{C}_\varepsilon^{\text{full}} = \text{Cov}(A_j, A_k)$$

are obtained by averaging against the stationary ε-dependent distribution

$$\mathbf{B}_\varepsilon^{\text{full}} = \begin{pmatrix} & & & 0 & 0 \\ & \mathbf{B}_\varepsilon & & 0 & 0 \\ & & & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{E14}$$

$$\mathbf{C}_\varepsilon^{\text{full}} = \begin{bmatrix} & & & \text{Cov}\left(\overline{z}, \chi_{[0,\varepsilon]}\right) & \text{Cov}\left(\overline{z}, \chi_{[1-\varepsilon,1]}\right) \\ & \mathbf{C}_\varepsilon & & \text{Cov}\left(U_\varepsilon, \chi_{[0,\varepsilon]}\right) & \text{Cov}\left(U_\varepsilon, \chi_{[1-\varepsilon,1]}\right) \\ & & & \text{Cov}\left(V_\varepsilon, \chi_{[0,\varepsilon]}\right) & \text{Cov}\left(V_\varepsilon, \chi_{[1-\varepsilon,1]}\right) \\ * & * & * & Q(1-Q) & -PQ \\ * & * & * & -PQ & P(1-P) \end{bmatrix} \tag{E15}$$

with $\mathbf{B}_\varepsilon$ and $\mathbf{C}_\varepsilon$ are defined as $\mathbf{B}$ and $\mathbf{C}$ but in the context of the truncated stationary distribution and constraints and with asterisks denoting the symmetrical terms. The boundary masses entering into matrix $\mathbf{C}_\varepsilon$ are defined as

$$Q = \int_0^\varepsilon \psi_\varepsilon(p) dp , \quad P = \int_{1-\varepsilon}^1 \psi_\varepsilon(p) dp \tag{E16}$$

The matrix $\mathbf{B}_\varepsilon^{\text{full}}$ is padded by zeros because $\partial_p \chi_{[0,\varepsilon]}$ and $\partial_p \chi_{[1-\varepsilon,1]}$ vanish in the interior of the subdomains. By neglecting the boundary terms coming from the integration by parts in the preceding calculation, we effectively limit the transfer of mass between the bulk and the boundaries so that the boundary masses agree with the continuous stationary distribution.

Note that the stationary allele frequency distribution $\overline{\psi}_\varepsilon(\mathbf{p})$ depends on the allele frequencies but also on the effective forces $\boldsymbol{\alpha}^*$. Therefore, the matrices $\mathbf{B}_\varepsilon$ and $\mathbf{C}_\varepsilon$, as well as matrices $\mathbf{B}_\varepsilon^{\text{full}}$ and $\mathbf{C}_\varepsilon^{\text{full}}$, obtained by averaging over the microscopic states are still functions of effective forces as in the continuous DynMaxEnt method. However, because the averages require integration

through the interior domain $[\varepsilon, 1 - \varepsilon]$ and separately through the boundaries $[0, \varepsilon]$ and $[1 - \varepsilon, 1]$, the integrals can no longer be written using the special functions and have to be evaluated numerically.

So far we have derived an extended dynamical system for the Lagrange multipliers $\boldsymbol{\alpha}^*$. One of the key questions is whether the dynamics converge to the target state $\boldsymbol{\alpha}$. Because the system has form

$$\frac{\partial \boldsymbol{\alpha}^*}{\partial t} = -\mathbf{M}_\varepsilon^*(\boldsymbol{\alpha}^* - \boldsymbol{\alpha}) \tag{E17}$$

where $\mathbf{M}_\varepsilon^* = 1/2N(\mathbf{C}_\varepsilon^{\text{full}})^{-1}\mathbf{B}_\varepsilon^{\text{full}}$ (note that the standard form of the ODE has a negative sign in front of $\mathbf{M}_\varepsilon^*$), and convergence to the steady-state is captured by the sign of the eigenvalues of $\mathbf{M}_\varepsilon^*$. When they are positive (*i.e.*, $-\mathbf{M}_\varepsilon^*$ has negative eigenvalues), the target state is asymptotically stable. It is easy to see that both $\mathbf{B}_\varepsilon^{\text{full}}$ and $\mathbf{C}_\varepsilon^{\text{full}}$ are symmetrical and positive semidefinite ($\mathbf{C}_\varepsilon^{\text{full}}$ is a covariance matrix and $\mathbf{u}^t\mathbf{B}_\varepsilon^{\text{full}}\mathbf{u} > 0$ can be written as a square), and thus the matrix $\mathbf{M}_\varepsilon^*$ has nonnegative eigenvalues. Therefore, the dynamics of Lagrange multipliers should converge to the fixed point. This fixed point is precisely the target point $\boldsymbol{\alpha}$ unless some of the eigenvalues converge to zero—then the dynamics may get trapped at a different point of the phase space and be unable to continue all the way to the target state.

Moreover, because the dynamics of Lagrange multipliers are five dimensional, while the dynamics of observables, driven by the $\mathbf{B}$ matrix, have only three degrees of freedom (because the entries of $\mathbf{B}_\varepsilon^{\text{full}}$ corresponding to A_4 and A_5 are zero), the method is in principle underdetermined. This essentially means that Lagrange multipliers μ and κ and ν and ρ may follow strange dynamics that together nevertheless lead to well-approximated observables. This is why in the next step we need to employ the constraints on κ and ρ by slaving their dynamics with dynamics of mutation rates:

$$\kappa = -\frac{1}{2N}\log(4N\mu) , \qquad \rho = -\frac{1}{2N}\log(4N\nu) \tag{E18}$$

Imposing these constraints leads to reduced dynamics with only three degrees of freedom, $\mathbf{A} = (\bar{z}, U_\varepsilon, V_\varepsilon)$ and $\boldsymbol{\alpha} = (\beta, \mu, \nu)$.

The full dynamics can be reduced by disregarding the entries of $\mathbf{B}_\varepsilon^{\text{full}}$ that are identically equal to zero and considering only the submatrix $\mathbf{B}_\varepsilon$. Intuitively, one expects that the reduced dynamics will be identical to the dynamics (14) governed by 3×3 matrices $\mathbf{B}_\varepsilon$ and $\mathbf{C}_\varepsilon = \text{Cov}(A_j, A_k)$. However, we show that the $\mathbf{C}$ matrix involves additional terms. To do that, we start by taking the first three equations from (E11), leaving out the equations for the dynamics of ρ and κ and by substituting ρ and κ with the expressions for μ and ν from (E18):

$$2N\sum_{k=1}^{3}\frac{d\alpha_k^*}{dt}\psi[A_k - \langle A_k\rangle] - \sum_{k=2}^{3}\frac{1}{\alpha_k^*}\frac{d\alpha_k^*}{dt}\psi[A_{k+2} - \langle A_{k+2}\rangle] + \partial_t\delta = -\partial_p\left[\frac{\sigma(p)}{2}\sum_{k=1}^{3}\left((\alpha - \alpha^*)\frac{\partial A_k}{\partial p}\right)\right] - \partial_p J[\delta, p]$$

We are left with a system of three ODEs that involve terms from the dynamics of κ and ρ (the second sum on the left). We then substitute values of κ and ρ from (E18) and obtain the correct $\mathbf{C}$ matrix:

$$\mathbf{C}_\varepsilon = \begin{bmatrix} \text{Var}(\bar{z}) & \text{Cov}(\bar{z}, U_\varepsilon) & \text{Cov}(\bar{z}, V_\varepsilon) \\ \text{Cov}(\bar{z}, U_\varepsilon) & \text{Var}(U_\varepsilon) & \text{Cov}(U_\varepsilon, V_\varepsilon) \\ \text{Cov}(\bar{z}, V_\varepsilon) & \text{Cov}(U_\varepsilon, V_\varepsilon) & \text{Var}(V_\varepsilon) \end{bmatrix} - \begin{bmatrix} 0 & \frac{1}{2N\mu}\text{Cov}\left(\bar{z}, \chi_{[0,\varepsilon)}\right) & \frac{1}{2N\nu}\text{Cov}\left(\bar{z}, \chi_{[1-\varepsilon,1)}\right) \\ 0 & \frac{1}{2N\mu}\text{Cov}\left(U_\varepsilon, \chi_{[0,\varepsilon)}\right) & \frac{1}{2N\mu}\text{Cov}\left(U_\varepsilon, \chi_{[1-\varepsilon,1)}\right) \\ 0 & \frac{1}{2N\mu}\text{Cov}\left(V_\varepsilon, \chi_{[0,\varepsilon)}\right) & \frac{1}{2N\nu}\text{Cov}\left(V_\varepsilon, \chi_{[1-\varepsilon,1)}\right) \end{bmatrix} \tag{E19}$$

resulting in a reduced dynamics $\partial\alpha^*/\partial t = (1/2N)(\mathbf{C}_\varepsilon)^{-1}\mathbf{B}_\varepsilon(\boldsymbol{\alpha} - \boldsymbol{\alpha}^*)$. Note that despite the original intuition, the added terms in the second and third columns must be present to reflect the dynamical character of the prefactors in the stationary distribution. The case with dominance is treated in an analogous way and results in matrices $\mathbf{B}_\varepsilon$ and $\mathbf{C}_\varepsilon$ of dimension 4×4.

An important question is whether the dynamics described in the preceding system converge to the state $\boldsymbol{\alpha}$. This is true for the dynamics of observables, driven by matrix $\mathbf{B}_\varepsilon$, no matter the initial and target state because of the symmetry and positive definiteness. This means that the approximation can, in principle, work, but it remains to be seen if the dynamics of Lagrange multipliers also have good convergence properties. The matrix $\mathbf{C}_\varepsilon$ is no longer symmetrical, and thus $(\mathbf{C}_\varepsilon)^{-1}B_\varepsilon$ is not necessarily positive definite. However, all our simulations suggest that positive semidefiniteness holds, and the only way that the approximation may fail to converge to the fixed point α is when some of the eigenvalues go to zero. In such a case, the trajectory gets trapped at a different state. We have observed this behavior for $4N\mu \ll 1$, where one of the mutation rates approaches zero.
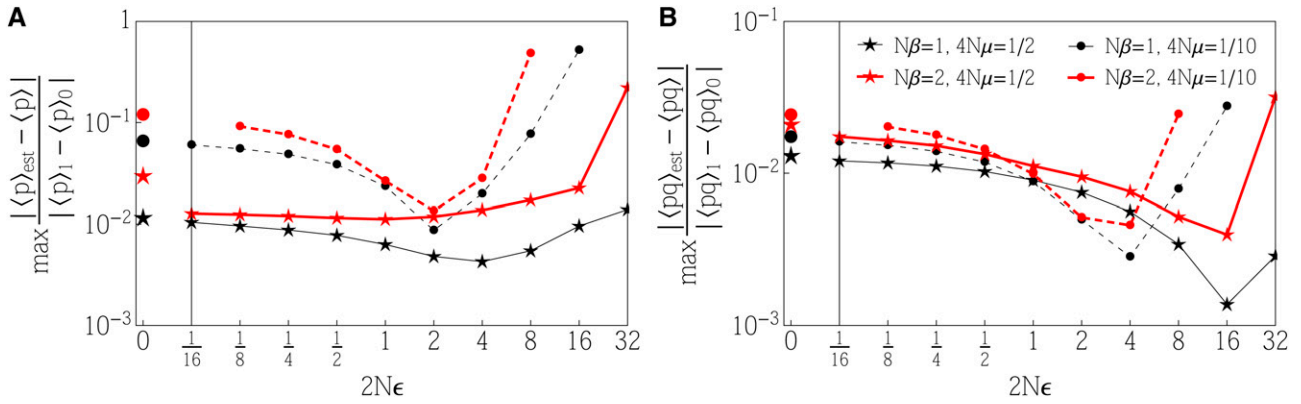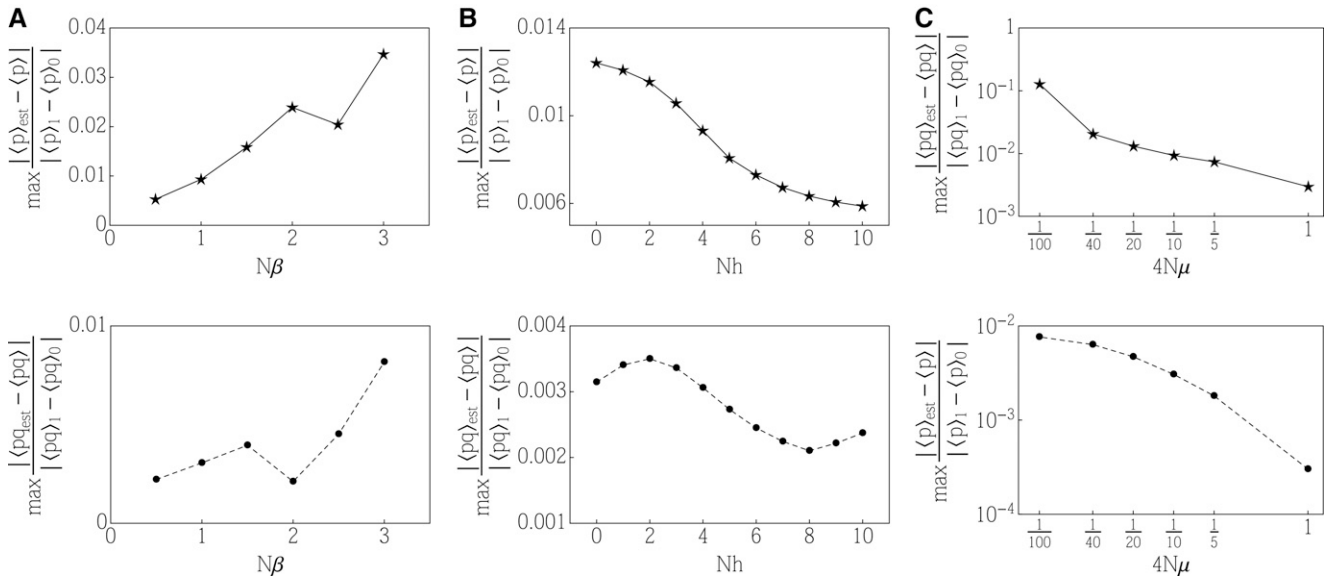
Figure F1 Effect of $N\varepsilon$ on performance of the general DynMaxEnt method. The response to a fast change in selection $N\beta = -1$ to 1 (or -2 to 2) has been compared with the approximated response of the system while keeping mutations fixed at $4N\mu = 4N\nu = 1/2$ (or $1/10$) and no overdominance $Nh = 0$ and $2N = 1000$. Each panel shows all four combinations of parameter regimes, distinguished by color ($N\beta$) and filling ($4N\mu$). We used a method that constrains expectations of three observables $\langle \bar{z} \rangle$, $\langle U \rangle$, and $\langle V \rangle$. Panels also contain the error of the continuous DynMaxEnt method plotted at $N\varepsilon = 0$ that was obtained by fixing effective mutations at $\mu^* = \mu$, $\nu^* = \nu$. (A) Relative error of the trait mean. (B) Relative error of the mean heterozygosity.

Appendix F: Dependence on the Truncation Parameter $N\varepsilon$

Evolution of quantitative traits depends on the following nondimensional parameters: $N\beta$, $N\mu$, and $N\nu$. The general DynMaxEnt approximation, as well as the piecewise-defined diffusion (E4), depends on an additional nondimensional parameter $N\varepsilon$ that influences the quality of the general DynMaxEnt approximation. As this parameter gets very small, the method approaches the continuous method without boundary layers. Because in the limiting case matrix $\mathbf{B}_\varepsilon \to \mathbf{B}$ contains diverging components, the error of the general approximation should increase as $N\varepsilon$ gets smaller. Similarly, the error will be large when $N\varepsilon \gg 1$. This is so because the relationship (22), which assumed $N\varepsilon \ll 1$, no longer matches the boundary masses of the approximated stationary distribution with the correct stationary distribution. Moreover, the boundary is too wide for $N\varepsilon \gg 1$ to disentangle the effects of selection and mutation.

Simulations in Figure F1 applied a sudden change in selection while the small mutation rates remained unperturbed. The figure shows that there exists an optimal value of the threshold $N\varepsilon$ that leads to the best approximation. This threshold seems to

Figure F2 Effect of $N\alpha$ on performance of the general DynMaxEnt method. The reference simulation undergoes a fast change in selection $N\beta : -1 \to 1$, heterozygosity $Nh : 4 \to 0$, and mutation $4N\mu, 4N\nu : 0.5 \to 0.1$ and $2N = 1000$. The three panels show dependence of the simulation error when the reference case parameters are changed one at a time. (A) Dependence on the strength of directional selection, changing from $-N\beta$ to $N\beta$ (plotted on a linear scale). (B) Dependence on the strength of selection for the heterozygous form, changing from Nh to 0 (plotted on a linear scale). (C) Dependence on the magnitude of mutation, changing from 0.5 to $4N\mu = 4N\nu$ (plotted on a logarithmic scale).
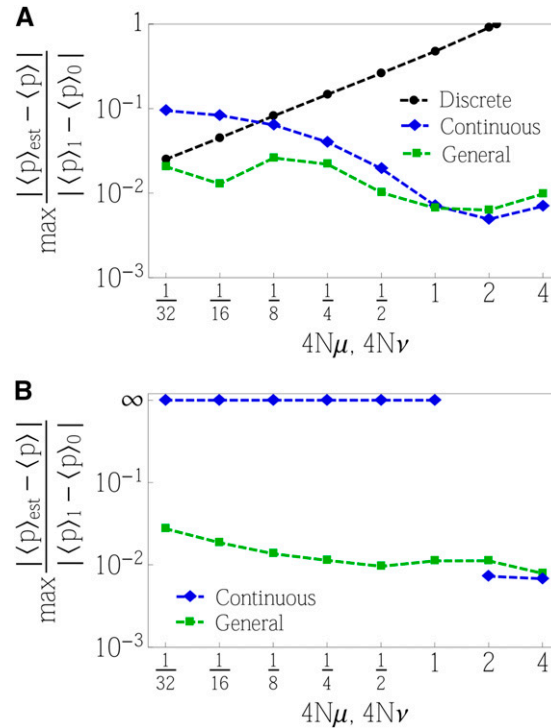
Figure G1 Comparison among discrete, continuous, and general approximations. We show performance of the methods in two scenarios: (A) rapid change in selection $N\beta = -1$ to 1 with the other parameters unperturbed $Nh = 0$, $N\mu = N\nu \in \{2^{-5}, \ldots, 2^2\}$; (B) rapid change in selection $N\beta = -1$ to 1 and mutation $4N\mu = 4N\nu = 1.1$ to $\{2^{-5}, \ldots, 2^2\}$ combined with a change in heterozygosity $Nh = 4$ to 0. The population size is $2N = 1000$.

depend on the mutation rates $N\mu, N\nu$ but also weakly on the selection rate $N\beta$, which is consistent with a representation of $p < N\varepsilon$ as a mutation-dominated regime. The optimal threshold differs when considering error in the trait mean and in the mean heterozygosity.

It is useful to know how the general DynMaxEnt method performs for realistic examples given a sensible fixed choice of $N\varepsilon$. We set $2N\varepsilon = 1$ such that ε corresponds to a frequency of one in the total of $2N$ individuals. We took the reference simulation in Example 2 and varied strengths of the parameters $N\beta$, Nh, and $4N\mu = 4N\nu$ one at a time. Figure F2 shows that the relative error in a trait mean is for most parameter settings within 1%. The worst-case scenario occurs when the selection gradient is strong and mutations drop rapidly to a small value. This is expected because strong and rapid change in selection implies a large deviation from adiabatic regime; thus the quasi-stationarity assumption may fail. Additionally, small mutations lead to a separation of timescales in the model that may be too complex to capture by a simple approximation. Surprisingly, stronger selection on heterozygosity does not decrease the accuracy of the method, which is counterintuitive because large Nh leads to more complex allele frequency distributions.

Appendix G: Comparison of Different Methods

We compare three available methods: the discrete approximation (see Appendix C), which captures responses to a change in selection in the limit $4N\mu, 4N\nu \to 0$; the continuous DynMaxEnt approximation (see Equation 14), valid for $4N\mu, 4N\nu > 1$ with a possibility to treat also small but static mutation; and the general DynMaxEnt approximation (see Equation 26), which extends the validity to dynamic mutations of an arbitrary magnitude. Note that dynamic Lagrange multipliers also may represent changing population size, not just the evolutionary forces.

Figure G1 shows two scenarios: perturbation of the system via a change of selection (A) and perturbation through changes in all evolutionary parameters (B). In the case of (A), the continuous DynMaxEnt method can be applied even in the case of small mutations $4N\mu < 1$ by enforcing a fixed mutation rate in time. But even then the continuous method performs worse than the general method, which shows a relative error on the order of a few percent. In comparison, the discrete approximation gives an accurate estimate when mutations drop below $4N\mu < 0.1$, but given its simplicity, it does not capture unequal and dynamic mutation rates.

The power of the general DynMaxEnt is demonstrated in Figure G1B for dynamic mutations. This situation cannot be captured by the simple discrete approximation, and the continuous DynMaxEnt, not valid for $4N\mu < 1$, shows divergence that is an
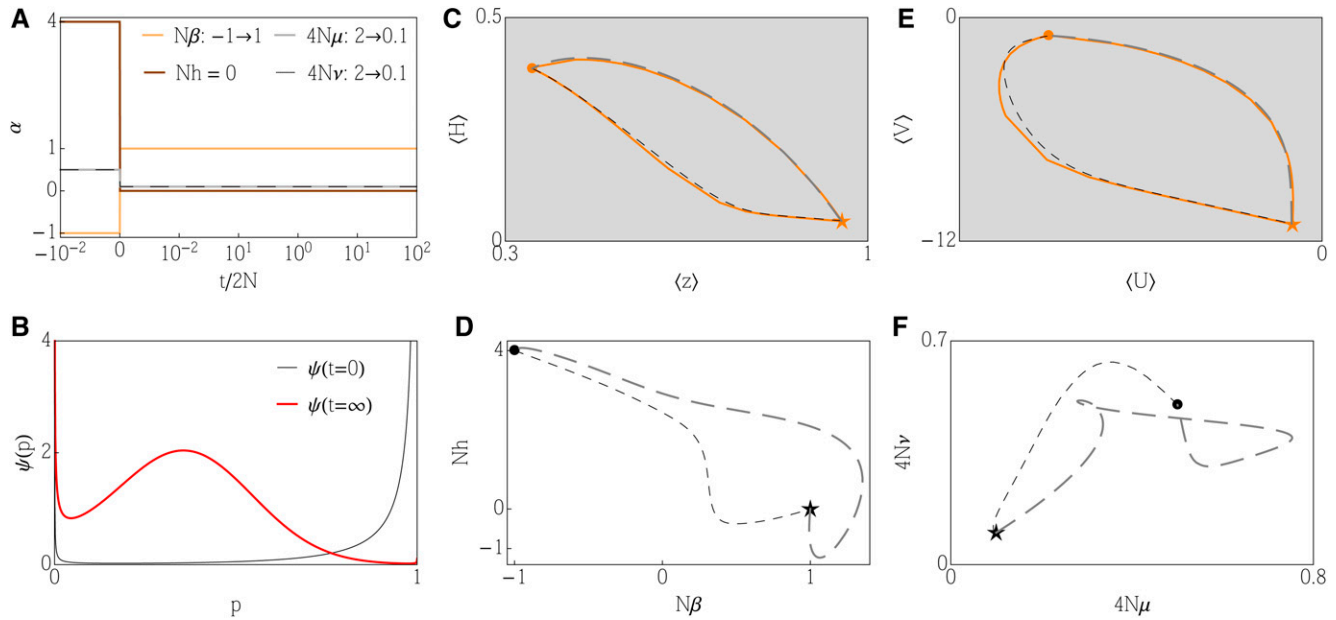
Figure H1 Irreversibility of the dynamics for Example 2. The dynamics in Figure 2 have been complemented by dynamics with the initial state and the final state reversed. (A) Values of evolutionary forces for the reversed simulation. (B) Initial and final allele frequency distributions for the reversed simulation. (C and E) The paths between the initial/final conditions (circle/star) displayed in the space of observables, together with the exact dynamics, shown in orange. We displayed the situation in two of six possible pairs of coordinates. (D and F) The dynamics visualized in the space of corresponding Lagrange multipliers. All figures show the forward trajectory (Example 2) in gray-dashed curve and dynamics with reversed initial and final state in black-dashed curve.

inherent property of the matrix **B**, approximating the dynamics of the observables. To show the performance of the continuous method, we included ∞ in the axis denoting the accuracy of the approximation.

## Appendix H: Irreversibility of the Adaptation Dynamics

Figure H1 compares the dynamics of the observables and Lagrange multipliers for Example 2, where the initially trimodal distribution loses polymorphism owing to a decrease in the selection for heterozygous individuals and a simultaneous decrease in the mutation rate and a reversed dynamics with the initial and final conditions switched. The dynamics are irreversible because switching the initial and final conditions leads to very different paths in the space of observables, as well as in the space of the effective forces. The dynamics in the space of observables show a good agreement with the exact solution of the problem. The paths from the initial to the final state are relatively straight. On the contrary, the dynamics of Lagrange multipliers show a complicated response of the system that exhibits a separation of timescales and overshooting of the equilibrium levels; *e.g.*, the effective selection coefficient grows to a level $N\beta \approx 1.5$ before decreasing back to $N\beta = 1$.

Appendix I: Equivalence Between Single-Locus Approximation and Multiple Loci with Equal Effects

A trait typically depends on many loci with different effects. If these effects are called γ_i, the simplest case is an additive trait $\bar{z} = \sum_i \gamma_i (p_i - q_i)$. If linkage equilibrium is assumed, the constraints can be written in terms of allele frequencies at the loci where each constraint is additive via loci ($A_i = \sum_k^n A_i^{(k)}$) with symmetrical expressions in terms of p_i. While the selection-related constraints depend on the distribution of effects γ linearly, the mutation-related constraints do not depend on it. If the effects of the L loci on an additive trait are the same, and $\gamma_i = 1$, then

$$\mathbf{B}_L = \left\langle \sum_{k=1}^{L} \frac{\partial \mathbf{A}}{\partial p_k} \frac{p_k q_k}{2} \frac{\partial \mathbf{A}^T}{\partial p_k} \right\rangle = \sum_{k=1}^{L} \left\langle \frac{\partial \mathbf{A}}{\partial p_1} \frac{p_1 q_1}{2} \frac{\partial \mathbf{A}^T}{\partial p_1} \right\rangle = L\mathbf{B}_1$$

where $\mathbf{B}_L$ is the matrix of genetic covariances for L loci. This is true even in the approximation for small dynamic mutations because the boundary terms are also additive across loci. A similar relationship holds for the covariance of fluctuations $\mathbf{C}_L = L\mathbf{C}_1$. Therefore, the matrix defining the MaxEnt dynamics of Lagrange multipliers in case of L loci with equal effects is identical to dynamics of Lagrange multipliers in case of a single locus, *i.e.*, $\mathbf{C}_L^{-1}\mathbf{B}_L = \mathbf{C}_1^{-1}\mathbf{B}_1$. As a result, the trait mean
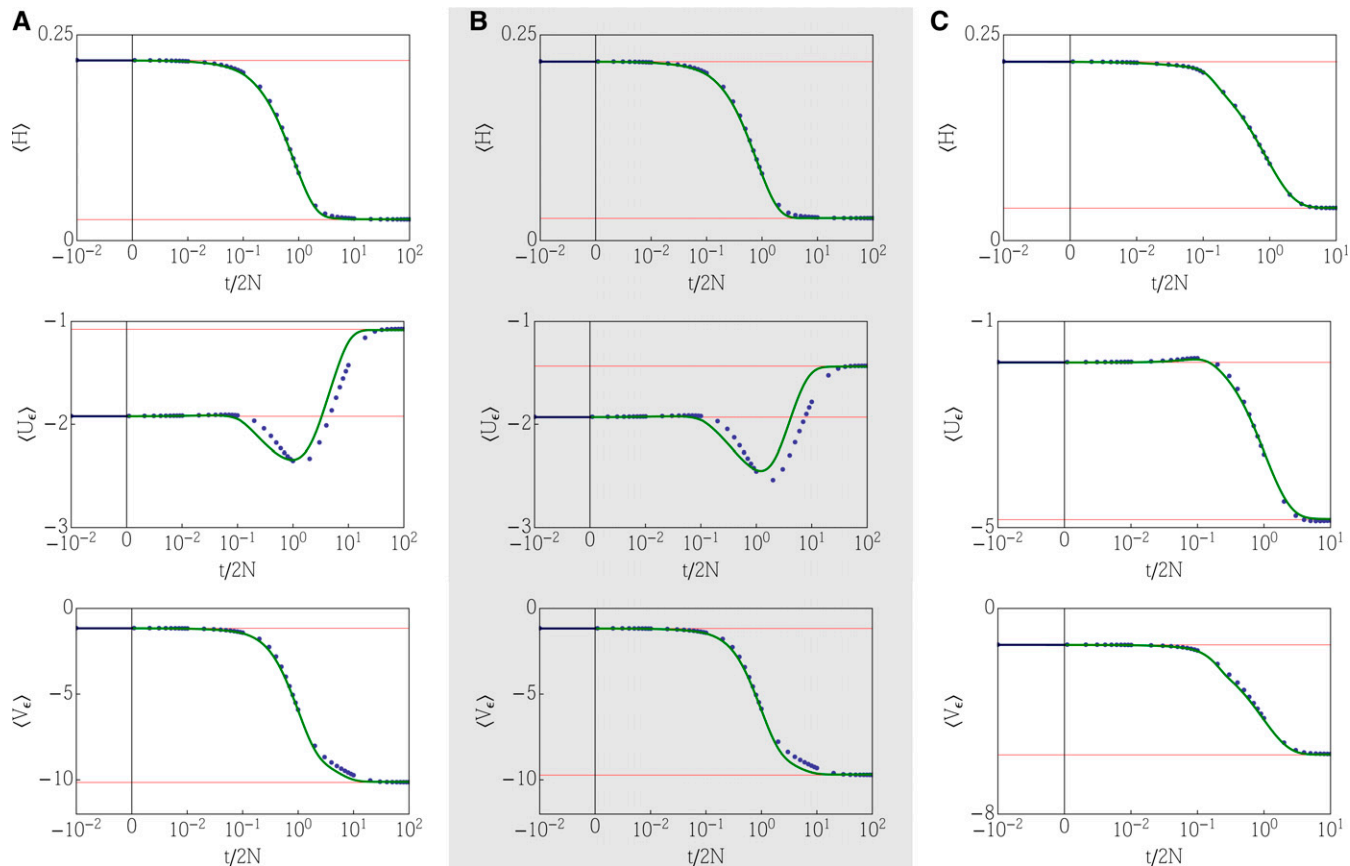
Figure J1 Example 3: dynamics of the quantitative traits for 100 loci of different effects. The effects are randomly drawn from (A) a uniform distribution in [0, 2], (B) an exponential distribution with mean 1, and (C) deterministic effects where 95 of the loci have effects 0.01, while the remaining 5 loci have effects 19.81. The true forces **a** change at time $t = 0$ and draw the system out of equilibrium. The response of the observed quantities $\langle pq \rangle$, $\langle 2\log p \rangle$, and $\langle 2\log q \rangle$ is shown in blue dots, while the approximation is shown in green. The quality of the approximation for the trait mean $\langle \bar{z} \rangle$ is shown in Figure 3.

$\bar{z}_n = \sum_{k=1}^{L}(p_i - q_i) = L(p - q) = n\bar{z}_1$, where p represents the allele frequency in the single-locus simulation with other parameters unchanged. However, the calculation of a trait mean for L independent loci of equal effects $\gamma_i = 1$ using the continuous model gives

$$\frac{d\bar{z}_L}{dt} = 2\sum_{k=1}^{L}\frac{dp_i}{dt} = L\frac{d\bar{z}_1}{dt} \tag{I.1}$$

showing that the relationship $\bar{z}_L = L\bar{z}_1$ follows also from the model description.

Appendix J: Additional Results for Multiple Loci with Unequal Effects

Figure J1 shows details on the quality of the general DynMaxEnt approximation in Example 3, where a linear trait depending on 100 loci with different effects evolves owing to a combination of a directional selection and selection on heterozygosity, random drift, and mutation. The changes in evolutionary forces are the same as in Example 2.

GENETICS

A General Approximation for the Dynamics of Quantitative Traits

Katarína Bod'ová, Gašper Tkačik, and Nicholas H. Barton

File S1: Mathematica notebook with the code for the general DynMaxEnt method. (.zip, 531 KB)

Available for download as a .zip file at:

http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.115.184127/-/DC1/FileS1.zip