When can fitness epistasis be ignored in a polygenic trait at equilibrium?

Archana Devi[a,*,1], Kavita Jain[b,2]

[a]Department of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090, Vienna, Austria
[b]Theoretical Sciences Unit, Jawaharlal Nehru Centre for Advanced Scientific Research,, Bangalore, 560064 , India

ARTICLE INFO

Keywords:
polygenic trait
fitness epistasis
equilibrium
stabilizing selection

ABSTRACT

Although many phenotypic traits are determined by a large number of genetic variants, the behavior of allele frequencies in a polygenic trait is not completely understood. The problem is especially challenging when the quantitative trait of interest is under epistatic selection as the allele frequency at a locus is affected by those at other loci. Here, we consider a panmictic, diploid finite population evolving under stabilizing selection and symmetric mutations when the population is in linkage equilibrium. In the stationary state, using a diffusion theory, we calculate the marginal distribution of allele frequency, and find parameter regimes where fitness epistasis can not be ignored for an accurate description of the frequency distribution. For such parameters, the mean deviation in the phenotypic optimum and genic variance are, however, found to be well captured even when epistatic interactions are neglected. Thus, while the presence of epistasis may not be evident in phenotypic quantities, it can strongly affect the allele frequency distribution. We also find that the allele frequency distribution at a locus is unimodal if its effect size is below a threshold effect and bimodal otherwise; these results are the stochastic analog of the deterministic ones where the stable allele frequency becomes bistable when the effect size exceeds a threshold. Our analytical results are verified against Monte Carlo simulations and numerical integration of a Langevin equation.

1. Introduction

The evolutionary dynamics of complex phenotypic traits have traditionally been studied in the framework of quantitative genetics that deals with the summary statistics of a trait; however, a detailed understanding of the genetic basis of phenotypic variation is crucial for predicting the evolutionary outcomes of quantitative traits such as complex diseases (García-González and O'Reilly, 2026). Population genetics focuses on the genetic details of the phenotypic trait as the allele frequencies underlying a trait evolve under the joint action of evolutionary forces such as mutation, selection, random genetic drift and gene flow, and informs us on the evolution and maintenance of phenotypic variation in a population and between populations and species. Due to the advancement of genome-wide association studies (GWAS) in recent years, much information about the genetic architecture of phenotypic traits has increasingly become available, and it is now known that the phenotypic variation depends on the number of genetic variants underlying a phenotype, the effect size and frequency of these variants, the interaction between these genetic variants and with the environment, etc. (Visscher, Wray, Zhang, Sklar, McCarthy, Brown and Yang, 2017; Timpson, Greenwood, Soranzo, Lawson and Richards, 2018). For these reasons, much work in the last two decades has focused on understanding quantitative traits at a genetic level and connecting these insights to the phenotype (Rockman, 2012; Jain and Stephan, 2017a; Sella and Barton, 2019; Barghi, Hermisson and Schlötterer, 2020).

The evolution of a quantitative trait in terms of the underlying allele frequencies is also theoretically challenging (Bürger, 2000; Walsh and Lynch, 2018) since stochastic models of polygenic adaptation are described by high-dimensional Fokker-Planck equations, and the phenotypic selection is typically epistatic, that is, it is nonlinear in the trait value and therefore can not be expressed as the sum of fitness of individual phenotypic values. Several recent studies have shown that phenotypic traits such as birth weight in humans are under stabilizing selection which has a tendency to reduce phenotypic variation in a population as the fitness of a trait value decreases with increasing deviation from the phenotypic optimum (Robertson, 1956; Kingsolver, Hoekstra, Hoekstra, Berrigan, Vignieri, Hill, Hoang,

archana.devi@univie.ac.at (A. Devi); jain@jncasr.ac.in (K. Jain)
https://ufind.univie.ac.at/en/person.html?id=1010614 (A. Devi); https://www.jncasr.ac.in/faculty/jain (K. Jain)
ORCID(s): 0000-0003-3597-7316 (A. Devi); 0000-0002-7406-4109 (K. Jain)

Gibert and Beerli, 2001; Sanjaka, Sidorenkoc, Robinson, Thornton and Visscher, 2018; de Villemereuil, Charmantier, Arlt, Bize, Brekke, Brouwer, Cockburn, Côté, Dobson, Evans, Festa-Bianchet, Gamelon, Hamel, Hegelbach, Jerstad, Kempenaers, Kruuk, Kumpula, Kvalnes, McAdam, McFarlane, Morrissey, Pärt, Pemberton, Qvarnström, Røstad, Schroeder, Senar, Sheldon, van de Pol, Visser, Wheelwright, Tufto and Chevin, 2020; Sodeland, Jentoft, Jorde, Mattingsdal, Albretsen, Kleiven, Wårøy Synnes, Espeland, Olsen, Andrè, Stenseth and Knutsen, 2022). If the population is initially far from the optimum, one can describe the evolutionary dynamics assuming directional selection which, however, lacks epistasis (Götsch and Bürger, 2024). But at large times, when the population is close to the optimum or has reached mutation-selection-drift equilibrium, the epistatic nature of stabilizing selection - both at phenotypic level as also at genetic level due to phenotype-genotype map - comes into play, and it is important to understand if and how these epistatic interactions affect the allele frequency distribution and the phenotypic quantities.

Here, we consider a polygenic trait evolving under stabilizing selection and symmetric mutations in a panmictic finite population under linkage equilibrium, and focus on its equilibrium properties. Due to stabilizing selection, the allele frequencies interact with each other and evolve such that the trait value approaches the optimum. Although the joint distribution of allele frequencies in the stationary state is known (Wright, 1937; Kimura, 1964), the distribution is not transparent enough to give an insight into the genetic architecture of the phenotypic trait. In this study, we obtain the marginal distribution of allele frequencies and study its properties in detail. Our main conclusion is that for a polygenic trait controlled by a large number of loci, as is relevant in many biologically realistic scenarios, the allele frequency distribution can be accurately described ignoring epistatic interactions, provided selection is sufficiently strong. But for weak to moderate selection, this is possible under certain conditions on mutation and selection parameters and if the locus under consideration has sufficiently small effect size. However, irrespective of selection strength, the genic variance is found to match well with Bulmer's expression (Bulmer, 1972) that has been obtained on neglecting epistasis. Our work also extends the previous theory for infinitely large populations (de Vladar and Barton, 2014; Jain and Stephan, 2017b) to large but finite populations.

2. Model

We consider the classic Latter-Bulmer model (Latter, 1960; Bulmer, 1972) that describes a panmictic, diploid population of size N characterized by a single polygenic trait which is controlled by L diallelic loci. The phenotype-genotype map is assumed to be additive (Hill, Goddard and Visscher, 2008), and the trait value $z = \sum_{i=1}^{L} \frac{\gamma_i}{2}(\sigma_i + \sigma_i^*)$ where σ_i, σ_i^*, respectively, denote the maternal and paternal contribution to the trait of an individual and take the value $1(-1)$ if mutant (wildtype) allele is present at the ith locus. Here the effect size of either allele is γ_i, and is chosen independently of the effect size at other loci from a common distribution $p(\gamma)$ which is assumed to be an exponential distribution with mean $\bar{\gamma}$ (Mackay, 2004; Goddard and Hayes, 2009).

The phenotypic trait evolves under stabilizing selection (Robertson, 1956; Kingsolver et al., 2001; Sanjaka et al., 2018; de Villemereuil et al., 2020; Sodeland et al., 2022) with the phenotypic fitness function,

$$w(z) \approx 1 - \frac{s}{2}(z - z_o)^2 \tag{1}$$

which decays quadratically away from a (time-independent) optimal trait value z_o, and where s denotes the strength of selection. We also assume that the mutation between wildtype and mutant allele occurs at an equal rate μ at each locus.

The dynamics follow a classical Wright-Fisher model where genotypes of the individuals undergo random genetic drift, selection, recombination, mutations, and random matings with other individuals to form the next generation - while selection, recombination, and mutation are deterministic forces, the allele frequencies fluctuate in a finite population due to random genetic drift. At large times, the population always reaches a steady state under the joint action of these biological forces, and here we focus on the resulting equilibrium state. Under linkage equilibrium (see Appendix A1), the genotype frequencies can be expressed as the product of allele frequency at each locus in a fully recombining population; however, the allele frequencies do not evolve independently due to the indirect epistatic interaction imposed by the fitness function (1).

In the following sections, the model described above is studied analytically using a diffusion theory, and numerically using Monte Carlo (MC) simulations and by integrating a Langevin equation using Euler-Maruyama (EM) method. As discussed in Sec. 3 and Sec. 4.4, for a population in linkage equilibrium, the MC simulations are accurate for low mutation rates, but EM method works better for high mutation rates. For this reason, below we describe both methods.

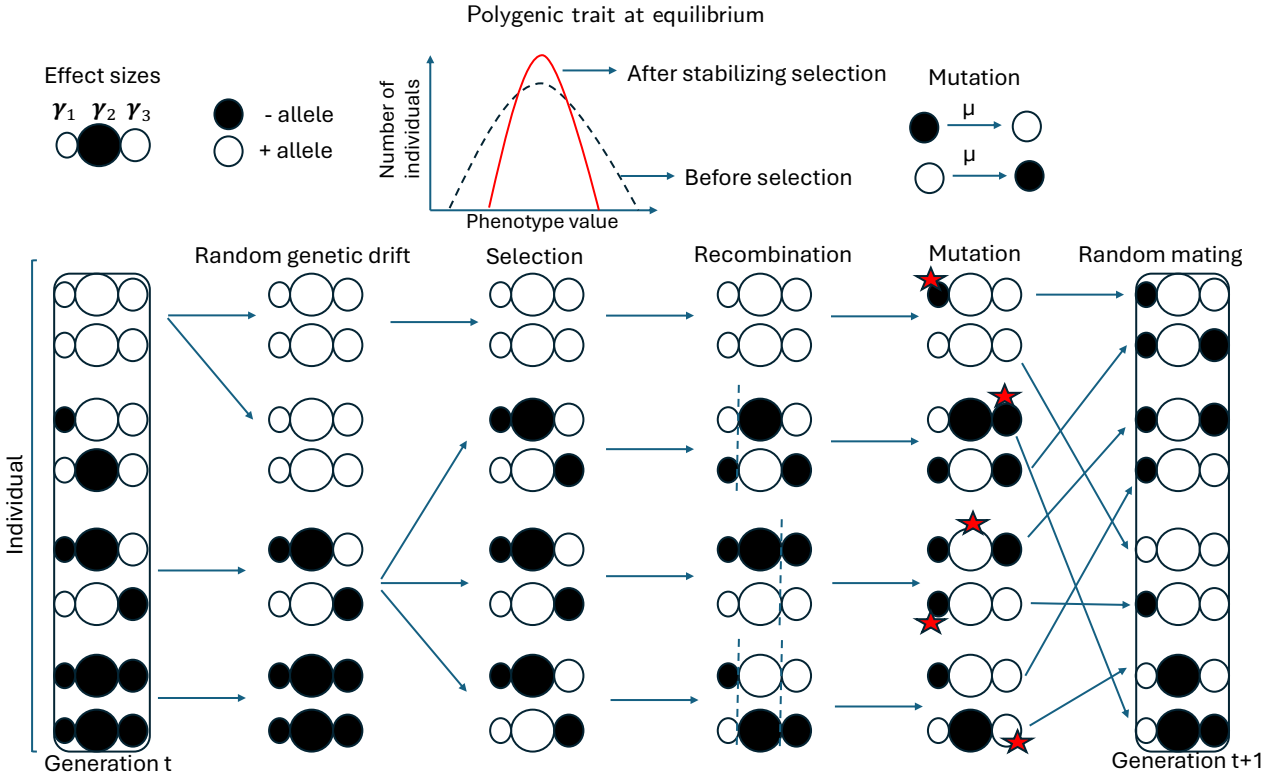

Figure 1: Diagram depicting how each evolutionary force shapes the genotypic configuration of individuals in a Wright-Fisher process implemented in Monte Carlo simulations. Here, we consider $N = 4$ diploid individuals with $L = 3$ loci to show the simulation steps from generation t to $t + 1$.

3. Methods

The model dynamics implemented in MC simulations are depicted in Fig. 1. We consider a population of N individuals, each having two chromosomes with L sites and assign an effect value γ to each locus (the effect sizes are exponentially distributed unless stated otherwise). Then each locus of a chromosome is multiplied by $+1$ or -1 with its effect size, where $+1$ or -1 are chosen with equal probability using a uniform random number generator. The trait value of each individual is half of the sum of its two chromosomal values. The population evolves in discrete and nonoverlapping generations through the following steps that are performed at each generation:

(i) Random genetic drift: We choose individuals randomly from the population using a uniform random number generator. However, the viability of a particular individual depends on its relative phenotypic fitness keeping the total number of individuals constant every generation.

(ii) Selection: Each individual is assigned a fitness according to (1). We weight the individual fitness with the maximum fitness in the population since the viability of the individual depends on its relative fitness in the population rather than its absolute fitness. We select the individual if a uniform random number is less than its relative fitness.

(iii) Recombination: Two chromosomes of the selected individual recombines according to the recombination parameter. Since we assume free recombination in this study, we find the recombination break points by drawing a Bernoulli-distributed random number with probability one half at every locus. After finding all the breakpoints, we exchange the genetic material of the two chromosomes at each recombination breakpoint. After the recombination event, we keep either of the recombined chromosomes with equal probability and discard the other one for every individual. This process continues until we select $2N$ chromosomes.

(iv) Mutation: We change the sign of the effect size of each locus if a standard uniform random number is less than the mutation probability.

(v) Random mating: We get N offspring for the next generation by pairing chromosomes randomly from the $2N$ chromosomes.

As shown in Appendix A2, under linkage equilibrium, the above model can be described using a diffusion theory. Following Itô prescription, corresponding to the Fokker-Planck equation (A2.1), the Langevin equation for the allele frequency x_i at the ith locus is given by [refer to Sec. 4.3.5, Gardiner (1997)]

$$dx_i(t) = \langle \delta x_i \rangle dt + \sqrt{\langle (\delta x_i)^2 \rangle}\, dW_i(t) \tag{2}$$

where, as discussed in Appendix A2,

$$\langle \delta x_i \rangle = -\frac{s}{2}[2\gamma_i x_i(1-x_i)\Delta c_1 + \gamma_i^2 x_i(1-x_i)(1-2x_i)] + \mu(1-2x_i) \tag{3}$$

$$\langle (\delta x_i)^2 \rangle = \frac{x_i(1-x_i)}{2N} \tag{4}$$

and dW_i is the Wiener process. The Langevin equation (2) is integrated numerically using the EM method (Higham, 2001). Dividing the time t into $t/\delta t$ intervals of equal length δt, the allele frequency at the ith locus is updated as (Sec. 4.3.1, Gardiner (1997))

$$x_i(n+1) = x_i(n) + \langle \delta x_i \rangle \delta t + \sqrt{\langle (\delta x_i)^2 \rangle}\, \eta\ ,\ 0 \le n \le \frac{t}{\delta t} \tag{5}$$

where η is a random variable chosen independently at each time step from a normal distribution with mean zero and variance δt. In all the numerical data presented using the above equation, the population starts with allele frequency one half at every locus, and time step size $\delta t = 0.1$.

For either method, after a burn-in period of $10N$ generations, the data in the stationary state are obtained by averaging over $10^5 - 10^8$ generations, depending on the parameters. Although the EM method is computationally faster and easier, as shown in Fig. 6, it fails to accurately capture the genic variance (see Appendix A1 and Sec. 4.4) for low mutation rates. The EM method is designed to simulate continuously fluctuating, large-population, and can not correctly model the small population dynamics or low mutation rate where fixation or loss of alleles occur (Hutzenthaler, Jentzen and Kloeden, 2011). On the other hand, the MC simulations which are computationally expensive for large populations with large number of loci, are accurate for low mutation rate; however, the linkage equilibrium gets distorted when both mutation and selection are strong in a finite population which generates (and maintains) non-random associations between loci faster than the maximum recombination probability equal to one half can dissipate them. These non-random associations prevent the population from reaching the perfect linkage equilibrium in MC simulations, but the EM method uses the Langevin equation which is derived assuming linkage equilibrium. For this reason, we use the appropriate method to obtain the numerical results discussed below.

4. Steady state distribution

In the steady state and under linkage equilibrium, the exact joint distribution of mutant allele frequencies has been obtained in the framework of diffusion theory, and is given by (Kimura, 1964) [also see Appendix A2]

$$P(\vec{x}) \propto e^{-Ns(c_1 - z_o)^2} \prod_{i=1}^{L} \psi_B(x_i) \tag{6}$$

where,

$$c_1(\vec{x}) = \sum_{i=1}^{L} \gamma_i (2x_i - 1) \tag{7}$$

is the population-averaged phenotype, and

$$\psi_B(x_i) \propto (x_i(1-x_i))^{4N\mu - 1} e^{-2Ns\gamma_i^2 x_i(1-x_i)} \tag{8}$$

Note that due to linkage equilibrium, although the population dynamics at each locus occur independently of the other loci, the distribution (6) is not a product measure; this is because due to epistasis, the selection (1) is nonlocal and depends on all loci through the global variable $c_1(\vec{x})$.

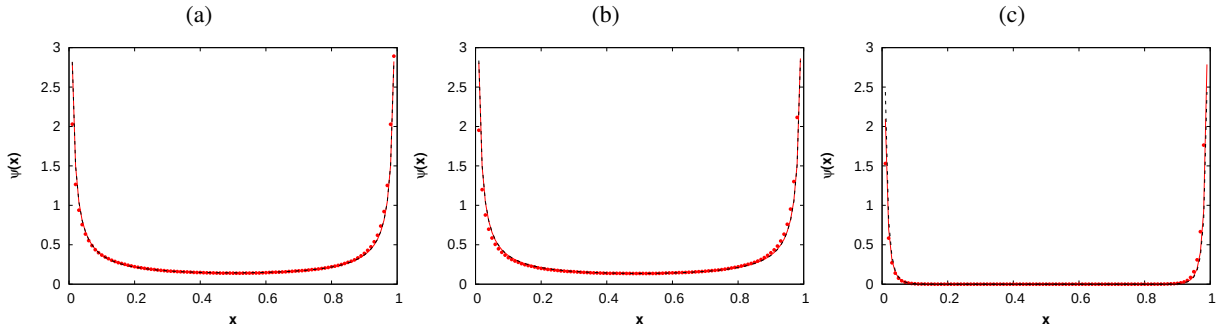

Figure 2: Marginal distribution for weak mutation ($4N\mu < 1$) at a locus with effect size (a) $\gamma_i = 0.01$, (b) $\gamma_i = 0.05$ and (c) $\gamma_i = 0.7$. The parameters are $N = 1000$, $s = 0.05$, $\mu = 0.00002$, $L = 200$, $\bar{\gamma} = 0.1$, and $z_o = 1$. The deterministic threshold size $\hat{\gamma} = 0.056$ for these parameters whereas stochastically, there is no threshold effect. The points are obtained from MC simulations and the red solid line shows the analytical expression (9) where $\kappa_2 \approx 4.8$ for the effect sizes used in this plot. Here, $2Ns\bar{\gamma}^2 = 1$ and as both κ_2 and $2Ns\kappa_2$ are large, (11) is satisfied and the marginal distribution (8) shown by black dashed line matches the distribution ψ.

4.1. Marginal distribution for a large number of loci

To understand the allele frequency distribution at a locus, we now consider the marginal distribution $\psi(x_i)$ of the frequency at the ith locus which can be obtained by integrating the joint distribution $P(\vec{x})$ over all the allele frequencies except x_i and using the central limit theorem. For a large number of loci, as shown in Appendix A3, the single-locus distribution can be approximated by

$$\psi(x_i) \propto (x_i(1-x_i))^{4N\mu-1} e^{-2Ns\gamma_i^2 x_i(1-x_i)} e^{\frac{-Ns[\gamma_i(2x_i-1)-z_o]^2}{1+2Ns\kappa_{2,i}}} \tag{9}$$

where,

$$\kappa_{2,i} = \sum_{j\neq i} 4\gamma_j^2(\langle x_j^2\rangle_{\psi_B} - \langle x_j\rangle_{\psi_B}^2) \tag{10}$$

is the *statistical variance* of allele frequency x_i w.r.t. the distribution ψ_B and given by (A3.6). As each summand on the right-hand side (RHS) of (10) is positive, $\kappa_{2,i}$ increases linearly with L and captures the effect of epistatic interactions, as explained below.

For $2Ns\kappa_{2,i} \gg 1$, the last factor in (9) reduces to $e^{-\frac{[\gamma_i(2x_i-1)-z_o]^2}{2\kappa_{2,i}}}$, and tends to unity provided $\kappa_{2,i}$ is also large (and z_o does not scale linearly with L). Thus, when $\kappa_{2,i} \gg 1$ *and* $2Ns\kappa_{2,i} \gg 1$, the effect of epistatic interactions can be ignored and the marginal distribution (9) reduces to $\psi_B(x_i)$:

$$\psi(x_i) \xrightarrow{\kappa_{2,i}\gg 1, 2Ns\kappa_{2,i}\gg 1} \psi_B(x_i) \tag{11}$$

On approximating $\kappa_{2,i}$ by $\kappa_2 = \sum_{j=1}^{L} 4\gamma_j^2(\langle x_j^2\rangle_{\psi_B} - \langle x_j\rangle_{\psi_B}^2)$, as shown in Appendix A4, we obtain

$$\kappa_2 \approx \begin{cases} \dfrac{2L\bar{\gamma}^2}{1+8N\mu} & , Ns\bar{\gamma}^2 \ll 2 \quad \text{(12a)} \\ 2L\bar{\gamma}^2 & , Ns\bar{\gamma}^2 \gg 2 \quad \text{(12b)} \end{cases}$$

For $\bar{\gamma} \sim 1$, when selection is strong ($2Ns\bar{\gamma}^2 \gg 1$), it is sufficient to ensure that $\frac{\kappa_2}{\bar{\gamma}^2}$ is also large in order to approximate the marginal distribution by ψ_B. Then (12b) shows that the epistatic factor in the marginal distribution can be ignored for large number of loci. However, for weak selection, κ_2 given by (12a) must be much larger than $(2Ns)^{-1}$ in order that epistasis can be neglected.

For large L, the exponential on the RHS of (9) can be expanded in powers of L^{-1}; however, as explained in Blinnikov and Moessner (1998), corrections to central limit theorem are required to obtain the correct expression to leading order in L^{-1}, and is briefly described in Appendix A5.

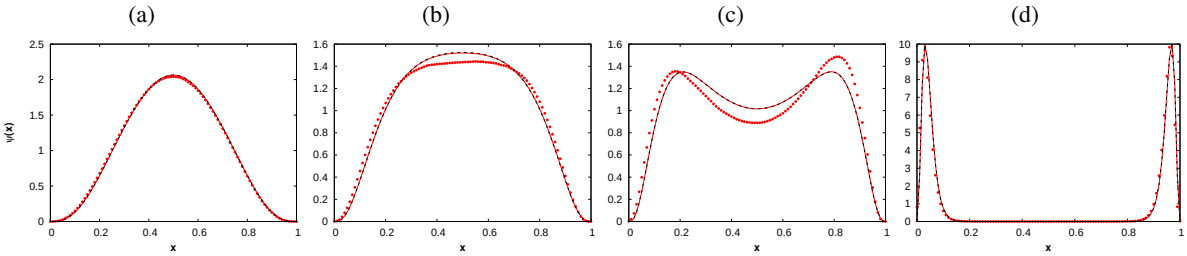

Figure 3: Marginal distribution for strong mutation ($4N\mu > 1$) and strong selection ($Ns\bar{\gamma}^2 > 2$) at a locus with effect size (a) smaller ($\gamma_i = 0.1$), (b) just below ($\gamma_i = 0.23$), (c) just above ($\gamma_i = 0.3$) and (d) larger ($\gamma_i = 0.7$) than the threshold size $\hat{\gamma}_N(L) = 0.24$. The parameters are $N = 1000$, $s = 0.1$, $\mu = 0.001$, $L = 200$, $\bar{\gamma} = 0.7$, and $z_o = 1$. For the set of effects used here, there were 137 large-effect loci. The points are obtained by solving (5) numerically, and the red solid line shows the analytical expression (9) where $\kappa_2 \approx 235$ for the effect sizes used in this plot. Here, as both κ_2 and $2Ns\kappa_2$ are large, (11) is satisfied and the marginal distribution (8) shown by black dashed line matches the distribution ψ.

4.2. Transition in marginal distribution for strong mutation

In Fig. 2, we show the marginal distributions obtained in simulations and in Fig. 3 and Fig. 4 by solving the Langevin equation (5) numerically, and compare them with the analytical expression (9). For weak mutation ($4N\mu < 1$), as Fig. 2 shows, the marginal distribution (9) is U-shaped, as expected. For the parameters in this figure, (12a) yields $\kappa_2 \approx 3.4, 2Ns\kappa_2 \approx 344$ so that (11) is satisfied, and we observe a good match with $\psi_B(x_i)$.

For strong mutation ($4N\mu > 1$), the distribution $\psi(x_i)$ has the following interesting property: it is unimodal if the effect size is below a threshold effect size $\hat{\gamma}_N(L)$ (see Sec. 4.5) and bimodal otherwise. This is displayed in Fig. 3 where the frequency distribution at a locus with sufficiently small effect size has one peak and it is flat for intermediate frequencies at a locus with the threshold size (inflection point), while at a locus with large effect size, the distribution has two maxima which are quite symmetric about frequency one half for parameters in Fig. 3 and asymmetric in Fig. 4. These behavior are also reminiscent of second order phase transitions in physical or chemical systems where the potential energy (given here by $-\ln\psi$) changes as an external parameter such as temperature (here, $2Ns\gamma_i^2$) is varied. For the parameters in Fig. 3, as $Ns\bar{\gamma}^2 = 49$, using (12b), we find that $\kappa_2 = 196$ which is comparable to value obtained using the effects used in this figure, and as selection is strong and κ_2 is large, we find that $\psi \approx \psi_B$; on the other hand, for the parameters in Fig. 4, since $Ns\bar{\gamma}^2 = 0.64$, (12a) yields $\kappa_2 = 0.75$ and as discussed above, the marginal distribution ψ is poorly estimated by ψ_B.

In Fig. 3 and Fig. 4, we also note a mismatch between the numerical data and the analytical expression (9) when the locus has an effect size close to or above the threshold size. This is because the population spends a long time near one of the maxima before traversing the valley separating the other maximum (Barton and Rouhani, 1987). Due to such shifts in the equilibria for large-effect loci, the numerical results for the stationary state distribution were obtained by averaging over 10^3 independent initial conditions (ensemble-averaging) as well as long time periods in the stationary state (time-averaging), as this allowed us to sample the distribution near both the allele frequency peaks efficiently.

4.3. Phenotypic mean deviation

Using the joint distribution (6), we find that the distribution of phenotypic mean (7) is given by

$$Pr(c_1) \propto e^{-Ns(\Delta c_1)^2} \int_0^1 \dots \int_0^1 \delta(c_1 - \sum_{i=1}^{L} \gamma_i(2x_i - 1)) \prod_{i=1}^{L} dx_i \, \psi_B(x_i) \tag{13}$$

$$\propto e^{-Ns(\Delta c_1)^2} e^{-\frac{c_1^2}{2\kappa_2}} \tag{14}$$

where $\Delta c_1 = c_1 - z_o$, κ_2 is given by (A4.2) and the last expression is obtained using the central limit theorem for large L (Bulmer, 1972; Lande, 1976). The above distribution gives the average and variance of the phenotypic mean to be

$$\langle c_1 \rangle = \frac{2Ns\kappa_2 z_o}{1 + 2Ns\kappa_2} \approx z_o \left(1 - \frac{1}{2Ns\kappa_2} + \mathcal{O}(L^{-2})\right) \tag{15}$$

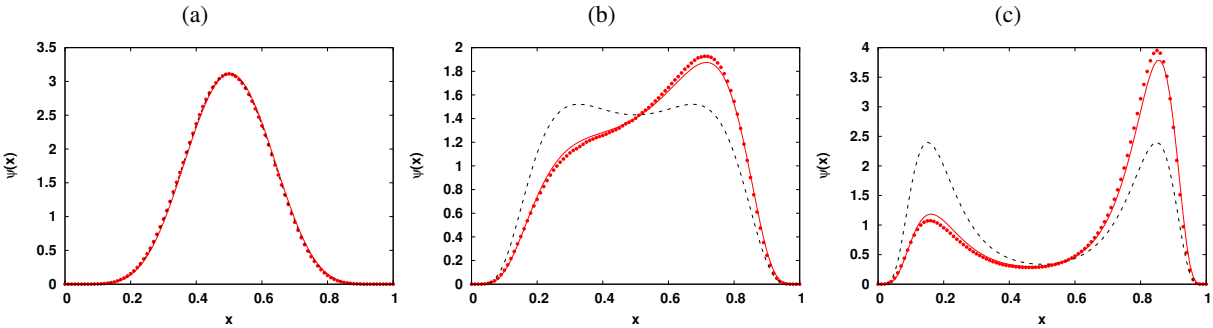

Figure 4: Marginal distribution for strong mutation ($4N\mu > 1$) and weak selection ($Ns\bar{\gamma}^2 < 2$) at a locus with effect size (a) smaller ($\gamma_i = 0.05$), (b) close to ($\gamma_i = 0.4$) and (c) larger ($\gamma_i = 0.5$) than the threshold size $\hat{\gamma}_N(L) \approx 0.42$. The parameters are $N = 1000$, $s = 0.1$, $\mu = 0.002$, $L = 1000$, $\bar{\gamma} = 0.08$, and $z_o = 2$. For the set of effects used here, there were 7 large-effect loci. The points are obtained by solving (5) numerically, and the red solid line shows the analytical expression (9) where $\kappa_2 \approx 1.24$ for the effect sizes used in this plot. Here, although $2Ns\kappa_2 \approx 248$ is large, as κ_2 is not very large, (11) is not satisfied and the marginal distribution (8) shown by black dashed line does not match the distribution ψ at loci except when the effect size is small.

$$\langle (c_1 - \langle c_1 \rangle)^2 \rangle \quad = \quad \frac{\kappa_2}{1 + 2Ns\kappa_2} \approx \frac{1}{2Ns} - \frac{1}{4N^2 s^2 \kappa_2} + \mathcal{O}(L^{-2}) \tag{16}$$

Then the deviation in the trait mean can be written as

$$\langle \Delta c_1 \rangle \approx -\frac{z_o}{2Ns\kappa_2} \tag{17}$$

From (15) and (12), we find that if the magnitude of the phenotypic optimum does not increase with the number of loci, the deviation in the mean phenotype vanishes with increasing L and the population is perfectly adapted (on an average). Intuitively, as the variance of mean becomes smaller with increasing L, the distribution of phenotypic mean gets narrower so that the first factor in (9) (although still random) can be neglected.

4.4. Mean genic variance

As discussed in Appendix A1, the genic variance within the population is defined as

$$c_2 = 2 \sum_{i=1}^{L} \gamma_i^2 x_i (1 - x_i) \tag{18}$$

where the mutant allele frequencies are random variables distributed according to (6). For equal effects and $z_o = 0$, Bulmer (1972) has obtained an expression for the expectation value of the genic variance by arguing as follows: from (6), the expected genic variance can be rewritten as

$$\langle c_2 \rangle \quad = \quad 2\gamma^2 \int_0^1 dx_1 ... dx_L \; Y e^{-4Ns\gamma^2 X^2} \hat{P}(\vec{x}) \tag{19}$$

where $X = \sum_{i=1}^{L} x_i - \frac{1}{2}, Y = \sum_{i=1}^{L} x_i(1 - x_i)$ and $\hat{P} \propto \prod_{i=1}^{L} [x_i(1 - x_i)]^{4N\mu - 1} e^{-2Ns\gamma_i^2 x_i(1-x_i)}$. Then, since X is antisymmetric about $x_i = 1/2$ while Y and $\hat{P}$ are symmetric, it follows that the expectation value of XY with respect to $\hat{P}$ is zero and it is concluded that "Hence X and Y must also be uncorrelated" (Bulmer, 1972). This reasoning then gives

$$\langle c_2 \rangle = \langle c_2 \rangle_{\psi_B} = 2\gamma^2 \int_0^1 dx_1 ... dx_L \; Y \hat{P}(\vec{x}) = 2L\gamma^2 \frac{2\mu N \; {}_1F_1\left(\frac{1}{2}; 4\mu N + \frac{3}{2}; \frac{1}{2}\gamma^2 Ns\right)}{(8\mu N + 1) \; {}_1F_1\left(\frac{1}{2}; 4\mu N + \frac{1}{2}; \frac{1}{2}\gamma^2 Ns\right)} \tag{20}$$

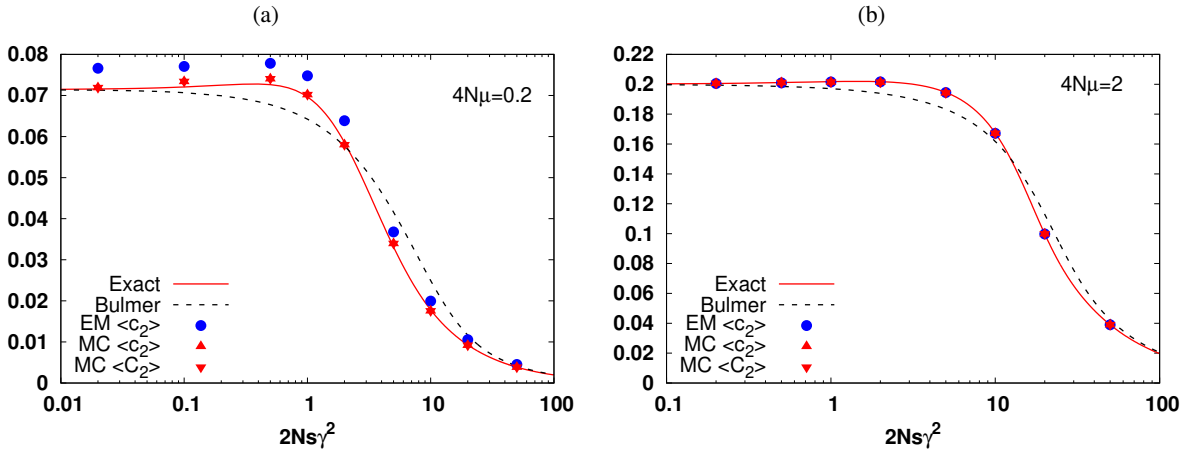

Figure 5: Mean genic variance (18) for $L = 2$ for (a) weak and (b) strong mutation. In all the figures, $N = 1000, z_o = 0, \gamma_i = 0.5$ (equal effects). In the legend, the line denoted by 'Exact' is obtained by numerically integrating the double integral of the joint distribution in the result (19) of diffusion theory, whereas 'Bulmer' shows the result (20) obtained by Bulmer, while 'MC' and 'EM', respectively, represent the data obtained using Monte-Carlo simulations and numerical integration of Langevin equation via Euler-Maruyama method.

where ${}_1F_1(a, b, z)$ is the confluent hypergeometric function [Chapter 13, (Olver, Olde Daalhuis, Lozier, Schneider, Boisvert, Clark, Miller, Saunders, Cohl and McClain, 2024)]. However, the exponential factor in the integrand on the RHS of (19) is an even function of X, and on expanding the exponential in a power series in X, we immediately find that the expectation value of $Ye^{-4Ns\gamma^2X^2}$ and Y w.r.t. $\hat{P}$ are not equal, contrary to Bulmer's argument above.

For two loci, the integrals in (19) can be readily done numerically and the expected variance is shown in Fig. 5a and Fig. 5b, and compared with $\langle c_2 \rangle_{\psi_B}$. We find that Bulmer's expression (20) does not match the exact result except for very weak selection (neutral limit) where the fitness epistasis is negligible and for very strong selection where variance in mean deviation given by (16) is negligible so that the loci evolve independently. These results are also compared with those obtained numerically. We find that the mean genic variance $\langle c_2 \rangle$ and mean genetic variance $\langle C_2 \rangle$ (where C_2 is defined in (A1.3)) obtained in MC simulations are equal since the linkage disequilibrium is negligible for weak to moderate mutation rate with small number of loci where the recombination probability of 0.5 can maintain the linkage equilibrium. The results from MC simulations are also in agreement with the exact results for both weak and strong mutation; however, the corresponding results for genic variance $\langle c_2 \rangle$ from EM method do not match the exact results for weak mutation for the reasons discussed in Sec. 3.

In Fig. 6 where genic variance and genetic variance are shown for large L, we find that the mean genic variance is close to $\langle c_2 \rangle_{\psi_B}$. However, the mean genic variance and the mean genetic variance in MC simulations are not equal for strong selection and strong mutation since the interference between genotypes is strong in this regime leading to a negative linkage disequilibrium between loci, and recombination with probability one half can not break all the associations between the loci; this results in a reduction in the genetic variance (Bulmer (1971, 1974)) which depends on the genic variance $\langle c_2 \rangle$ and the fitness variance $1/s$ (Turelli (1988); Turelli and Barton (1990, 1994)). Figure 6 also shows that the EM method does not work for weak mutations, but it is in closer agreement with Bulmer's expression than the MC simulation data when mutation is strong. From (9), we find that for large L, the corrections to Bulmer's result are of order L^{-1}, and these are discussed in Appendix A6.

4.5. Threshold effect size

We finally describe the threshold effect size introduced in Sec. 4.2, and how it depends on L and N. For $L \to \infty$, taking the derivative of the distribution (8) w.r.t. x_i, we immediately find that for $4N\mu > 1$, the allele frequency distribution has one maximum if the effect size at a locus is below the threshold effect size $\hat{\gamma}_N$ and two maxima otherwise, where $\hat{\gamma}_N$ is given by

$$\hat{\gamma}_N = \sqrt{\frac{8\mu}{s} - \frac{2}{Ns}} = \sqrt{\frac{2}{Ns}(4N\mu - 1)} \tag{21}$$

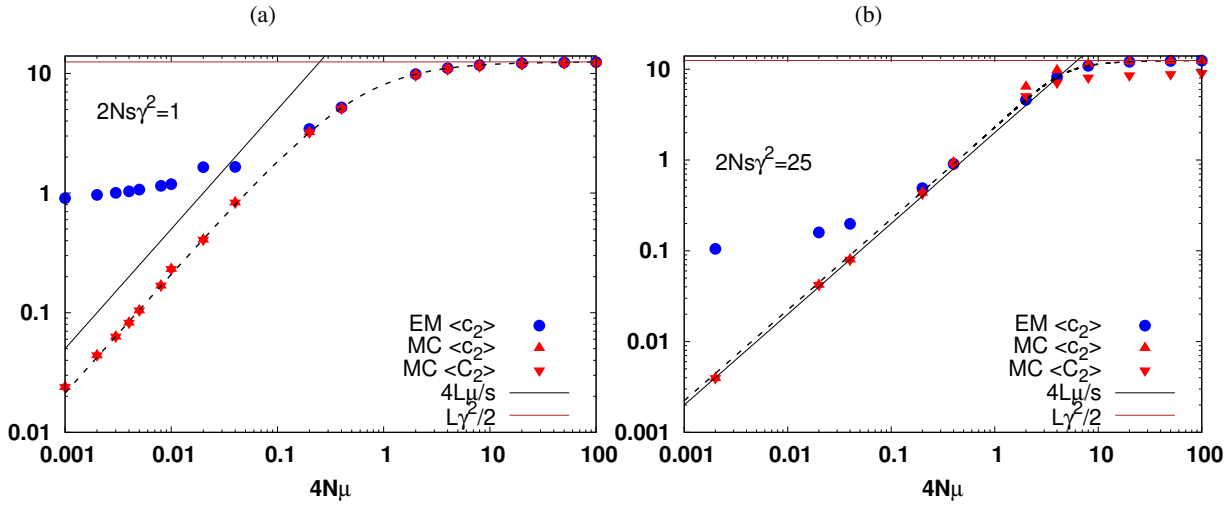

Figure 6: Mean genic variance (18) and mean genetic variance (A1.3) for $L = 100$ for (a) weak and (b) strong selection. In all the figures, $N = 1000, z_o = 0, \gamma_i = 0.5$ (equal effects). In the legend, the 'MC' and 'EM', respectively, represent the data obtained using Monte-Carlo simulations and numerical integration of Langevin equation via Euler-Maruyama method. The black dashed line shows the result (20) obtained by Bulmer, while the black solid line shows the genic variance in the rare allele approximation (Barton, 1986), and the brown line shows the genic variance corresponding to all frequencies equal to one half in the strong mutation limit.

The maxima of the allele frequency distribution (8) occur at

$$\chi_i = \begin{cases} \frac{1}{2} & , \gamma_i < \hat{\gamma}_N \quad (22a) \\ \frac{1}{2}\left(1 \pm \sqrt{1 - \frac{\hat{\gamma}_N^2}{\gamma_i^2}}\right) & , \gamma_i > \hat{\gamma}_N \quad (22b) \end{cases}$$

and for $\gamma_i > \hat{\gamma}_N$, the maxima of the bimodal distribution are separated by a minimum at frequency $1/2$. Thus, (21) and (22) show that when selection is weaker than mutation ($\frac{Ns\gamma_i^2}{2} \ll 4N\mu - 1$) or equivalently, for *small effect size* ($\gamma_i \ll \hat{\gamma}_N$), due to a high degree of polymorphism, the allele frequency distribution peaks at one half; however, when selection is stronger than mutation or for *large effect size*, the allele frequency distribution peaks at frequencies away from one half but the population is not monomorphic (unlike the U-shaped distribution in weak mutation regime).

The finite L corrections to $\hat{\gamma}_N$ and χ_i are derived in Appendix A7, and shown in Fig. 7; we find that the modes of the allele frequency distribution at a locus with effect size away from the threshold effect $\hat{\gamma}_N(L)$ are well approximated by the result (22) for infinite loci. But close to the threshold effect, they are substantially different; in particular, for positive (negative) phenotypic optimum, the maximum in the frequency distribution of the small-effect locus increases (decreases) with the effect size and occurs at a frequency which is substantially larger (smaller) than one half.

It is also useful to compare these results for an infinitely large population where the deterministic allele frequency p_i obeys (de Vladar and Barton, 2014; Jain and Stephan, 2017b)

$$\frac{dp_i}{dt} = \frac{-sp_i(1-p_i)}{2}\left[2\gamma_i(m_1 - z_o) + \gamma_i^2(1-2p_i)\right] + \mu(1-2p_i) \quad (23)$$

with the deterministic trait mean, $m_1 = \sum_{i=1}^{L} \gamma_i(2p_i - 1)$. Assuming that the trait mean deviation is zero in the stationary state, it has been shown that the deterministic allele frequency is in stable equilibrium below a threshold

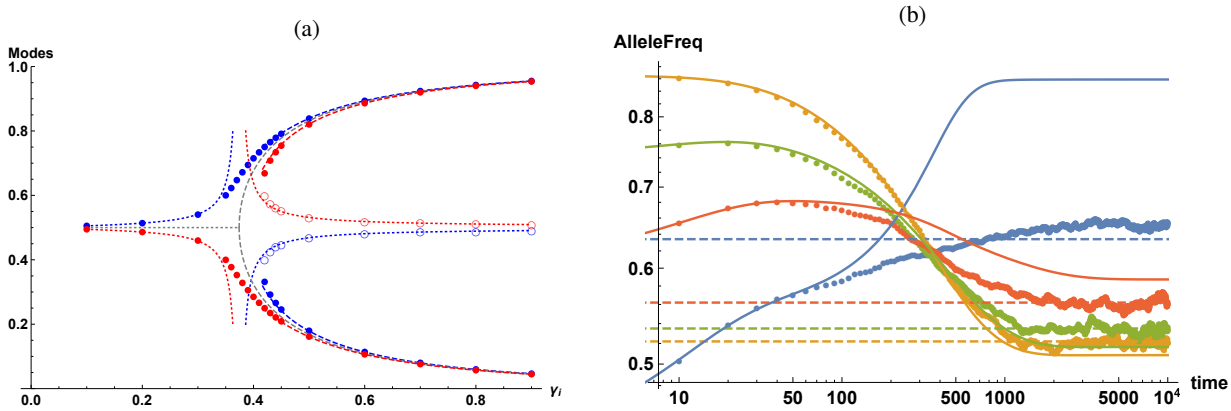

Figure 7: (a) Allele frequency at which the modes in the stationary state marginal allele frequency distribution occur for $z_o = 2$ (blue) and -2 (red). The points are obtained by numerically solving the cubic equation (A7.1) for the modes, with closed (open) symbols denoting the maximum (minimum). The dotted and dashed lines, respectively, show the approximate expressions in (A7.3a) and (A7.3b) for the mode frequency for finite number of loci while the gray lines show (22) for infinite number of loci. The parameters are the same as in Fig. 4, *viz.*, $N = 1000, s = 0.1, \mu = 0.002, \bar{\gamma} = 0.08$, and $L = 1000$. The threshold effect obtained from the cubic equation (A7.2) and the approximation (A7.5) are, respectively, 0.416 and 0.401, while (21) yields $\hat{\gamma}_N(L \to \infty) = 0.374$. (b) Dynamics of deterministic allele frequency (solid lines) obtained by numerically solving (23) and average allele frequency (points) obtained by numerically solving (5) for 5000 independent stochastic runs for effect size 0.24 (yellow), 0.36 (green), 0.52 (red), 0.8 (blue), keeping the initial frequencies to be the same in both deterministic and stochastic model. In either model, the population is initially equilibrated to an optimum at $z_o = 0$ and then the optimum is suddenly shifted to $z_o = 1$ in response to which the allele frequencies evolve in time. The other parameters are $L = 200, s = 0.05, \mu = 0.002, N = 200$. The threshold effect $\hat{\gamma}_N(L)$ is 0.359 and 0.493 when the optimum is at zero and one, respectively. The dashed lines show the average allele frequency $\langle x_i \rangle$ in the stationary state for phenotypic optimum equal to one, and are computed using the marginal distribution (9).

effect $\hat{\gamma} = \sqrt{8\mu/s}$ and is bistable above it (de Vladar and Barton, 2014), and given by

$$p_i^* = \begin{cases} \dfrac{1}{2} & , \gamma_i < \hat{\gamma} \quad (24a) \\ \dfrac{1}{2}\left(1 \pm \sqrt{1 - \dfrac{\hat{\gamma}^2}{\gamma_i^2}}\right) & , \gamma_i > \hat{\gamma} \quad (24b) \end{cases}$$

5. Discussion

In this article, we considered a model that describes a large but finite population under stabilizing selection and symmetric mutation in the equilibrium state. The dynamics of adaptation following a sudden change in the phenotypic optimum of such a model have been studied in deterministic (de Vladar and Barton, 2014; Jain and Stephan, 2015, 2017b) and stochastic setting (Thornton, 2019; Hayward and Sella, 2022; Höllinger, Wölfl and Hermisson, 2023).

Here we asked: in what parameter regimes the fitness epistasis can be neglected when the population is in equilibrium? Our key result for the allele frequency distribution is summarized in (11). We find that for strong selection, it is sufficient to have large number of loci in order to describe the distribution without taking epistasis into account, but for weak to moderate selection, there are conditions on mutation and selection parameters given by (12) under which epistasis can be neglected.

To connect these results at the genetic level to the phenotypic quantities, we also studied the phenotypic deviation and mean genic variance. We find that Bulmer's argument (Bulmer, 1972) for mean genic variance is incorrect and therefore, their expression given by (20) is not exact, but it is a good approximation when the number of loci are large. We stress that the disagreement with Bulmer's expression is not because of linkage disequilibrium [Bulmer effect; (Bulmer, 1971)] as the diffusion theory formulated here is in linkage equilibrium rather the point is that (20) neglects

the epistatic effects. The discussion in Bulmer (1972) does not invoke the number of loci, and we are not aware of other work where this error has been pointed out. Thus, for small number of loci, as in oligogenic traits (Bell, 2009; Boyle, Li and Pritchard, 2017), the expression (29) of Bulmer (1972) does not hold but it is a good approximation for large L for any selection strength.

Taken together, our results show that for weak to moderate selection, although genic variance matches well with Bulmer's expression, epistasis strongly affects the allele frequency distribution at loci with large effect size. Thus, while epistasis effects may not be evident at the phenotypic level, they can strongly influence the distribution at the genetic level. In a recent study by Courau, Lambert and Schertzer (2026), the model here has been generalized to include heterogeneous mutation rates and asymmetric mutations, and the authors arrive at the same general conclusions as in this article but via a different route.

We close the article by discussing the key differences in the stationary state of infinite (de Vladar and Barton, 2014; Jain and Stephan, 2017b) and finite populations: first, it should be noted that, in general, there is no one-to-one correspondence between the multistability in a deterministic model and multimodality in the corresponding stochastic model. This point, for example, has been made in the context of biochemical reaction systems where some monostable (bistable) deterministic systems have been observed to have a bimodal (unimodal) distribution when stochastic fluctuations are allowed (Hahl and Kremling, 2016). Here, as $\hat{\gamma} > \hat{\gamma}_N$, when $\hat{\gamma}_N < \gamma_i < \hat{\gamma}$, the deterministic allele frequencies are monostable but the allele frequency distribution is bimodal; however, outside this parameter range, the bistability (monostability) in the deterministic model corresponds to bimodality (unimodality) in the stochastic model.

Second, the stationary state solution of the Fokker-Planck equation (A2.1) is unique [refer to Chapter 5, (van Kampen, 1997)]; that is, it is independent of the initial allele frequencies for both small- and large-effect loci. Furthermore, as already mentioned, in a finite population, the allele frequency of a large-effect locus spends a long time (presumably, exponentially long in population size) near one of the maxima before shifting to the other maximum. In contrast, in an infinite population, the stationary state allele frequency of a large-effect locus depends on the initial condition and does not shift between the two solutions given in (24b). These points are further illustrated in Fig. 7b where the stationary state frequencies in the deterministic and stochastic model, starting from the same initial condition, are found to be quite close for small effect loci but not for the large-effect ones. For the latter case, to obtain a match between the $N \to \infty$ limit of the stochastic model and the deterministic model, one also need to average over the initial conditions in the deterministic model.

A1. Linkage equilibrium

Let f_{σ,σ^*} denote the frequency of a diploid individual with maternal and paternal sequences, σ and σ^*, respectively. If the phenotypic effect size at locus i is $\gamma_i/2$, for equivalent sexes and assuming additive phenotype-genotype map, the trait value of the individual can be written as

$$z = \sum_{i=1}^{L} \frac{\gamma_i}{2}(\sigma_i + \sigma_i^*) \tag{A1.1}$$

with $\sigma_i = 1(-1)$ for mutant (wildtype) allele. Under random mating, the parent frequencies factorize in Hardy-Weinberg equilibrium (HWE). Then the mean phenotype of the population (denoted by bar) obtained by averaging over the phenotypes of all individuals is given by

$$c_1 = 2\sum_{i=1}^{L} \frac{\gamma_i}{2}\overline{\sigma_i} = \sum_{i=1}^{L} \gamma_i(2x_i - 1) \tag{A1.2}$$

where $\overline{\sigma_i^*} = \overline{\sigma_i}$ for equivalent sexes, and x_i is the frequency of the mutant allele at locus i. Similarly, one obtains the within-population (additive) genetic variance as (Bulmer, 1971)

$$C_2 = \overline{z^2_{\sigma,\sigma^*}} - (\overline{z})^2 = 2\sum_{i\neq j} \frac{\gamma_i\gamma_j}{4}[\overline{\sigma_i\sigma_j} - \overline{\sigma_i}\,\overline{\sigma_j}] + 2\sum_{i} \frac{\gamma_i^2}{4}[\overline{\sigma_i^2} - \overline{\sigma_i}^2] \tag{A1.3}$$

Here, we assume that the population is in linkage equilibrium, $\overline{\sigma_i\sigma_j} = \overline{\sigma_i}\,\overline{\sigma_j}$, and obtain the (additive) genic variance, c_2 given by (18).

A2. Diffusion theory

Let $\vec{x} = \{x_1, ..., x_L\}$ denote the mutant allele frequency vector and $P(\vec{x}, t)$ is the joint distribution of these frequencies. The forward Fokker-Planck equation (FPE) is then given by

$$\frac{\partial P}{\partial t} = -\sum_{i=1}^{L} \frac{\partial}{\partial x_i}[\langle \delta x_i \rangle P] + \frac{1}{2}\sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j}[\langle \delta x_i \delta x_j \rangle P] \tag{A2.1}$$

where, the mean, variance and covariance in the change in the allele frequencies due to stabilizing selection, mutation and random genetic drift are described below.

First consider the effect of selection: Since $x_i = \sum_{\sigma,\sigma} \delta_{\sigma_i,1} f_{\sigma,\sigma^*}$, the (deterministic) change in the mutant allele frequency after selection is given by (Kimura, 1964; Bulmer, 1972)

$$\langle \delta x_i \rangle_s = -\frac{s}{2}[2\gamma_i x_i(1 - x_i)(\bar{z} - z_o) + \gamma_i^2 x_i(1 - x_i)(1 - 2x_i)] \tag{A2.2}$$

on using that the population is in HWE before selection. For symmetric mutation rates, the deterministic change in x_i due to mutation is given by

$$\langle \delta x_i \rangle_m = \mu(1 - x_i) - \mu x_i = \mu(1 - 2x_i) \tag{A2.3}$$

Thus the change $\langle \delta x_i \rangle$ due to selection and mutation is obtained by adding their respective contributions, and given by

$$\langle \delta x_i \rangle = -\frac{s}{2}[2\gamma_i x_i(1 - x_i)\Delta c_1 + \gamma_i^2 x_i(1 - x_i)(1 - 2x_i)] + \mu(1 - 2x_i) \tag{A2.4}$$

Note that $\langle \delta x_i \rangle$ depends on the frequency of other loci also through $\Delta c_1 = \sum_{j=1}^{L} \gamma_j(2x_j - 1) - z_o$. Finally, the change due to multinomial sampling in the Wright-Fisher process is given by

$$\langle \delta x_i \delta x_j \rangle = \frac{x_i(1 - x_i)}{2N}\delta_{ij} \tag{A2.5}$$

where the covariance vanishes as the population is in linkage equilibrium.

In the steady state [refer to Chapter 6, (Risken, 1996); (Kimura, 1964)], the LHS of (A2.1) is zero, and therefore, the divergence of the total probability current vanishes: $\sum_{i=1}^{L} \frac{\partial J_i(\vec{x})}{\partial x_i} = 0$, where

$$J_i = \langle \delta x_i \rangle P - \frac{1}{4N} \frac{\partial}{\partial x_i} [x_i(1-x_i)P] \tag{A2.6}$$

In general, it is not necessary that each J_i be same for all loci, and moreover, equal to zero. But if we do assume that $J_i = 0$ for all i, we obtain

$$\frac{\partial \ln P}{\partial x_i} = \frac{4N\langle \delta x_i \rangle - (1-2x_i)}{x_i(1-x_i)} \equiv F_i(\vec{x}) \tag{A2.7}$$

Then, for $\Phi = \ln P$, the total derivative $d\Phi = \sum_i \frac{\partial \Phi}{\partial x_i} dx_i$ exists iff $\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}$; using (A2.4), we verify that this condition is indeed satisfied, and therefore, we can write

$$\ln P(\vec{x}) = \sum_{i=1}^{L} \int dx_i \frac{4N\langle \delta x_i \rangle - (1-2x_i)}{x_i(1-x_i)} + e^C \tag{A2.8}$$

We thus obtain

$$P(\vec{x}) = C\, e^{-Ns(c_1-z_o)^2} e^{-2Ns\sum \gamma_i^2 x_i(1-x_i)} \prod_{i=1}^{L} [x_i(1-x_i)]^{4N\mu-1} \tag{A2.9}$$

where, the constant C is determined using the normalization condition, $\int_0^1 \prod_{i=1}^{L} dx_i\; P(\vec{x}) = 1$.

A3. Marginal distribution

The marginal distribution $\psi(x_i)$ of the allele frequency in the stationary state can be found by integrating the joint distribution $P(\vec{x})$ given by (6) over the frequencies of all but the ith locus; this gives

$$\psi(x_i) \propto \psi_B(x_i) \int_0^1 \ldots \int_0^1 e^{-Ns(c_1-z_o)^2} \prod_{j\neq i} dx_j \psi_B(x_j) \tag{A3.1}$$

where

$$\psi_B(x_i) \propto e^{-2Ns\gamma_i^2 x_i(1-x_i)} (x_i(1-x_i))^{4N\mu-1} \tag{A3.2}$$

and $c_1 = \sum_{j=1}^{L} \gamma_j(2x_j - 1)$. To evaluate the multiple integrals in (A3.1), we take $X' = \sum_{j\neq i} \gamma_j(2x_j - 1)$ and rewrite its RHS as

$$\frac{\psi(x_i)}{\psi_B(x_i)} \propto \int_{-\Gamma'}^{\Gamma'} dX' e^{-Ns(X'+\gamma_i(2x_i-1)-z_o)^2} \times \int_0^1 \ldots \int_0^1 \delta[X' - \sum_{j\neq i} \gamma_j(2x_j-1)] \prod_{j\neq i} dx_j \psi_B(x_j) \tag{A3.3}$$

where $\Gamma' = \sum_{j\neq i} \gamma_j$, and X' is a sum of independent but non-identically distributed random variables chosen from $\psi_B(x_j)$. The inner integrals on the RHS of the above equation can be calculated by appealing to the central limit theorem for large L, and we get

$$\frac{\psi(x_i)}{\psi_B(x_i)} \overset{L\gg 1}{\propto} \int_{-\infty}^{\infty} dX' e^{-Ns(X'+\gamma_i(2x_i-1)-z_o)^2} e^{-\frac{(X'-\kappa_{1,i})^2}{2\kappa_{2,i}}} \tag{A3.4}$$

$$\propto \exp\left[-\frac{Ns(\gamma_i(2x_i-1)-z_o)^2}{1+2Ns\kappa_{2,i}}\right] \tag{A3.5}$$

where $\kappa_{1,i}$ and $\kappa_{2,i}$ are, respectively, the mean and variance of the sum $X' = \sum_{j\neq i} \gamma_j(2x_j - 1)$ when averaged over the (normalized) distribution $\prod_{j\neq i} \psi_B(x_j)$.

Since $\psi_B(x_i)$ is symmetric about $x_i = 1/2$, it follows that the mean $\kappa_{1,i} = 0$. In the marginal distribution $\psi(x_i)$, the effect of other $L-1$ loci due to epistatic interactions in the phenotypic fitness function appears through the variance $\kappa_{2,i}$. Since $\langle x_i \rangle_{\psi_B} = 1/2$ for all loci, the variance $\kappa_{2,i} = \sum_{j\neq i} 4\gamma_j^2(\langle x_j^2\rangle_{\psi_B} - \langle x_j\rangle_{\psi_B}^2)$ which shows that $\kappa_{2,i}$ is simply a weighted sum of the variance of the distribution $\psi_B(x_i)$, and is given by

$$\kappa_{2,i} = \sum_{j=1, j\neq i}^{L} \frac{\gamma_j^2}{1+8N\mu} \frac{{}_1F_1(\frac{3}{2}, 4N\mu + \frac{3}{2}, \frac{Ns\gamma_j^2}{2})}{{}_1F_1(\frac{1}{2}, 4N\mu + \frac{1}{2}, \frac{Ns\gamma_j^2}{2})} \tag{A3.6}$$

where ${}_1F_1(a, b, z)$ is the confluent hypergeometric function.

A4. Statistical variance in the absence of epistasis

Here, we study the properties of $\kappa_2 = \sum_{j=1}^{L} 4\gamma_j^2(\langle x_j^2\rangle_{\psi_B} - \langle x_j\rangle_{\psi_B}^2)$. Note that κ_2 is the standard statistical variance while the mean genic variance $\langle c_2\rangle_{\psi_B}$ is related to the expectation of $x(1-x)$, although both are calculated w.r.t. ψ_B. We have

$$\kappa_2 = \sum_{j=1}^{L} 4\gamma_j^2 \left(\frac{1}{4(1+8N\mu)} \frac{{}_1F_1(\frac{3}{2}, 4N\mu + \frac{3}{2}, \frac{Ns\gamma_j^2}{2})}{{}_1F_1(\frac{1}{2}, 4N\mu + \frac{1}{2}, \frac{Ns\gamma_j^2}{2})} \right) \tag{A4.1}$$

$$\approx L \int_0^\infty \frac{p(\gamma)\gamma^2}{1+8N\mu} \frac{{}_1F_1(\frac{3}{2}, 4N\mu + \frac{3}{2}, \frac{Ns\gamma^2}{2})}{{}_1F_1(\frac{1}{2}, 4N\mu + \frac{1}{2}, \frac{Ns\gamma^2}{2})} \tag{A4.2}$$

where $p(\gamma)$ is the distribution of effects. For exponentially-distributed effect size with mean $\bar{\gamma}$, we have

$$\kappa_2 \approx \frac{L\bar{\gamma}^2}{1+8N\mu} \int_0^\infty \frac{{}_1F_1(\frac{3}{2}, 4N\mu + \frac{3}{2}, \frac{Ns\bar{\gamma}^2}{2}y^2)}{{}_1F_1(\frac{1}{2}, 4N\mu + \frac{1}{2}, \frac{Ns\bar{\gamma}^2}{2}y^2)} y^2 e^{-y} dy \tag{A4.3}$$

As the above integral does not appear to be exactly solvable, we estimate it by noting that the ratio of the confluent hypergeometric function in the above integrand is a monotonically increasing function of its argument, and for *fixed* $U = 4N\mu$, this ratio may be approximated using (13.7.1) and (13.11.1) of Olver et al. (2024) as

$$\frac{{}_1F_1(\frac{3}{2}, U + \frac{3}{2}, a)}{{}_1F_1(\frac{1}{2}, U + \frac{1}{2}, a)} \approx \begin{cases} 1 + \dfrac{4Ua}{4U^2 + 8U + 3}, & a \ll 1 \quad \text{(A4.4a)} \\ (2U+1)\left(1 - \dfrac{U}{a}\right), & a \gg 1 \quad \text{(A4.4b)} \end{cases}$$

Thus, for a given $4N\mu$, using (A4.4) in (A4.3), we obtain

$$\kappa_2 \approx \begin{cases} \dfrac{L\bar{\gamma}^2}{1+8N\mu}\left(2 + \dfrac{1536\bar{\gamma}^2 N^2\mu^2}{\hat{\gamma}^2(8N\mu+1)(8N\mu+3)}\right), & Ns\bar{\gamma}^2 \ll 2 \quad \text{(A4.5a)} \\ L\bar{\gamma}^2\left(2 - \dfrac{\hat{\gamma}^2}{\bar{\gamma}^2}\right), & Ns\bar{\gamma}^2 \gg 2 \quad \text{(A4.5b)} \end{cases}$$

where, $\hat{\gamma} = \sqrt{\frac{8\mu}{s}}$.

When selection is weak ($Ns\bar{\gamma}^2 \ll 2$) and mutation is strong ($4N\mu \gg 1$), (A4.5a) shows that κ_2 approaches zero as $1/N$; in contrast, when selection is stronger than mutation, due to (A4.5b), the variance $\kappa_2 \to 2L\bar{\gamma}^2$ so that it remains nonzero in the deterministic limit. These results can be understood as follows: as the width of $\psi_B(x_i)$ about a maximum is expected to decrease with N, for large N, one may approximate $\psi_B(x_i)$ by a Dirac delta function centred at $1/2$ for

small effect locus, and an average of two Dirac delta functions located at $p_{i,\pm} = \frac{1}{2}\left(1 \pm \sqrt{1 - \frac{\hat{\gamma}^2}{\gamma_i^2}}\right)$ for large-effect locus (de Vladar and Barton, 2014). Thus, for small-effect locus, as the distribution is unimodal and sharply-peaked in large populations, the variance vanishes in the deterministic limit. But for large-effect locus, as a consequence of the bimodality, the distribution remains broad resulting in a nonzero variance.

A5. Corrections to central limit theorem

Consider the distribution of the random variable, $Y = \frac{\sum_{j=1}^{L} \gamma_j(2x_j-1)}{\sqrt{\kappa_2}}$ where x_i's are independently distributed according to the (normalized) distribution, $\psi_B(x_i) \propto e^{-2Ns\gamma_i^2 x_i(1-x_i)}(x_i(1-x_i))^{4N\mu-1}$ and $\kappa_2 = \sum_{j=1}^{L} \langle \gamma_j^2(2x_j-1)^2\rangle_{\psi_B}$ where we have used that the mean $\langle \gamma_j(2x_j-1)\rangle_{\psi_B} = 0$. Then, from (43) of Blinnikov and Moessner (1998), we have

$$\mathrm{Prob}(Y) \propto e^{-\frac{Y^2}{2}}\left[1 + \frac{\kappa_4}{\kappa_2^2}\frac{3-6Y^2+Y^4}{24} + \mathcal{O}(L^{-2})\right] \tag{A5.1}$$

where, $\kappa_4 = \sum_{j=1}^{L} \langle \gamma_j^4(2x_j-1)^4\rangle_{\psi_B} - 3\langle \gamma_j^2(2x_j-1)^2\rangle_{\psi_B}^2$ is the fourth cumulant of the random variable $\gamma_j(2x_j-1)$ w.r.t. the (normalized) distribution $\psi_B(x_j)$. Using the above result in (A3.3) for $X' = \sqrt{\kappa_2}Y - \gamma_i(2x_i-1)$ and performing the integral over X', we arrive at

$$\frac{\psi(x_i)}{\psi_B(x_i)} \propto 1 + \frac{\kappa_4 - 4\kappa_2(\gamma_i(2x_i-1)-z_o)^2}{8\kappa_2^2} + \mathcal{O}(L^{-2}) \tag{A5.2}$$

A6. Stationary state genic variance

Consider the genic variance (18) when it is averaged over the stationary state distribution (A5.2). For $z_o \ll L$, we can write

$$\begin{aligned}
\langle x_i(1-x_i)\rangle &= \int_0^1 dx_i x_i(1-x_i)\psi(x_i) && \text{(A6.1)}\\
&\approx \frac{\int_0^1 dx_i x_i(1-x_i)\psi_B(x_i)\left(1 - \frac{(\gamma_i(2x_i-1)-z_o)^2}{2(\kappa_2+\frac{\kappa_4}{8\kappa_2})} + \ldots\right)}{\int_0^1 dx_i \psi_B(x_i)\left(1 - \frac{(\gamma_i(2x_i-1)-z_o)^2}{2(\kappa_2+\frac{\kappa_4}{8\kappa_2})} + \ldots\right)} && \text{(A6.2)}\\
&= \frac{\int_0^1 dx_i x_i(1-x_i)\psi_B(x_i)}{\int_0^1 dx_i \psi_B(x_i)}\\
&+ \frac{1}{2(\kappa_2+\frac{\kappa_4}{8\kappa_2})}\frac{\int_0^1 dx_i x_i(1-x_i)\psi_B(x_i)\int_0^1 dx_i\psi_B(x_i)(\gamma_i^2(2x_i-1)^2+z_o^2)}{[\int_0^1 dx_i\psi_B(x_i)]^2}\\
&- \frac{1}{2(\kappa_2+\frac{\kappa_4}{8\kappa_2})}\frac{\int_0^1 dx_i x_i(1-x_i)\psi_B(x_i)(\gamma_i^2(2x_i-1)^2+z_o^2)}{\int_0^1 dx_i\psi_B(x_i)} && \text{(A6.3)}
\end{aligned}$$

On summing over all the loci, the last equation gives the mean genic variance to be

$$\frac{\langle c_2\rangle}{L} = \frac{U}{1+2U}\int_0^\infty d\gamma\gamma^2 p(\gamma)\frac{\mathcal{F}(\frac{3}{2},\gamma)}{\mathcal{F}(\frac{1}{2},\gamma)}$$

$$+ \quad \frac{1}{(\kappa_2 + \frac{\kappa_4}{8\kappa_2})} \frac{U}{1+2U} \int_0^\infty d\gamma \gamma^4 p(\gamma) \left(\frac{U+1}{2U+3} \frac{\mathcal{F}(\frac{5}{2},\gamma)}{\mathcal{F}(\frac{1}{2},\gamma)} - \frac{U}{2U+1} \frac{\mathcal{F}^2(\frac{3}{2},\gamma)}{\mathcal{F}^2(\frac{1}{2},\gamma)} \right) \tag{A6.4}$$

where, for brevity, we have defined $\mathcal{F}(\alpha,\gamma) = {}_1F_1\left(\frac{1}{2}, 4N\mu + \alpha, \frac{Ns\gamma^2}{2}\right)$ and $U = 4N\mu$.

The first term on the RHS of (A6.4) is independent of z_o and is equal to $\langle c_2 \rangle_{\psi_B}$, and the second term that gives the L-dependent correction is of order L^{-1} and is also seen to be independent of the phenotypic optimum. Thus, when $z_o \ll L$, the mean genic variance (to $\mathcal{O}(L^{-1})$) does not depend on the location of the phenotypic optimum.

A7. Threshold effect size and mode frequency

On setting the derivative $\frac{d\psi}{dx_i}$ of the marginal distribution (9) equal to zero, we arrive at the following cubic equation for the modes $\chi_i(L)$:

$$(4N\mu - 1)\frac{(1-2\chi_i)}{\chi_i(1-\chi_i)} - 2Ns\gamma_i^2(1-2\chi_i) + \frac{4Ns\gamma_i(z_o - \gamma_i(2\chi_i - 1))}{1 + 2Ns\kappa_{2,i}} = 0 \tag{A7.1}$$

The above equation has two complex roots and one real root below the threshold effect $\hat{\gamma}_N(L)$ and three real roots above it. This change in the behavior of $\chi_i(L)$ occurs when the discriminant of the above cubic polynomial [(1.11.12), (Olver et al., 2024)] is equal to zero:

$$D = a_2^2 a_1^2 - 4a_3 a_1^3 + 18 a_0 a_2 a_3 a_1 - a_0\left(4a_2^3 + 27 a_0 a_3^2\right) = 0 \tag{A7.2}$$

where $a_i, i = 0,1,2,3$ are the coefficients of x^i in (A7.1). The resulting equation is a 3rd order equation in $\hat{\gamma}_N^2(L)$ and can be solved numerically to obtain the threshold effect for finite N and L.

One can, however, obtain an analytical expression for $\hat{\gamma}_N(L)$ when the number of loci are very large. Since κ_2 grows linearly with L, for infinite number of loci, the last term on the LHS of (A7.1) vanishes yielding (21) and (22) for the threshold effect $\hat{\gamma}_N$ and mode allele frequency χ_i, respectively, which are independent of z_o. For large L, using (A3.5), we find that the steady state distribution has maximum at

$$\chi_i(L) = \begin{cases} \frac{1}{2} + \frac{1}{2Ns\kappa_2} \frac{\gamma_i z_o}{\hat{\gamma}_N^2 - \gamma_i^2} & , \gamma_i < \hat{\gamma}_N(L) \quad \text{(A7.3a)} \\ \frac{1}{2}\left(1 \mp \sqrt{1 - \frac{\hat{\gamma}_N^2}{\gamma_i^2}}\right) + \frac{1}{2Ns\kappa_2} \frac{\hat{\gamma}_N^2 (z_o \pm \sqrt{\gamma_i^2 - \hat{\gamma}_N^2})}{2\gamma_i(\gamma_i^2 - \hat{\gamma}_N^2)} & , \gamma_i > \hat{\gamma}_N(L) \quad \text{(A7.3b)} \end{cases}$$

where $\hat{\gamma}_N(L)$ denotes the threshold effect for finite L. For $\gamma_i > \hat{\gamma}_N(L)$, the minimum in the allele frequency is given by the expression on the RHS of (A7.3a).

As shown in Fig. 7, a threshold effect exists below which (A7.1) has only one real root and corresponds to the maximum in the unimodal distribution. But, above the threshold effect, two additional real roots of (A7.1) appear which give the allele frequency at which the minimum and the second maximum of the bimodal distribution occur. Thus, at the threshold effect, for positive (negative) z_o, the minimum and the low-frequency (high-frequency) maximum of the bimodal distribution coincide. On matching (A7.3a) and the relevant solution in (A7.3b), we get

$$\pm 2Ns\kappa_2 \Delta^3 = 3\hat{\gamma}_N^2 z_o \pm \hat{\gamma}_N^2 \Delta + 2z_o \Delta^2 \tag{A7.4}$$

where $\Delta^2 = \hat{\gamma}_N^2(L) - \hat{\gamma}_N^2$ decreases with increasing L. The above cubic equation for Δ is exactly solvable, and has two complex roots and one real root. Here, we estimate the real root by noting that the first term on the RHS which is independent of L can be balanced if $\kappa_2 \Delta^3$ is also independent of L thus yielding

$$\hat{\gamma}_N(L) \approx \hat{\gamma}_N + \frac{1}{2\hat{\gamma}_N}\left(\frac{3\hat{\gamma}_N^2 |z_o|}{2Ns\kappa_2}\right)^{2/3} + \mathcal{O}(L^{-1}), \; z_o \neq 0 \tag{A7.5}$$

which shows that the deviation $\hat{\gamma}_N(L) - \hat{\gamma}_N$ decays rather slowly with L. The above expression also shows that the threshold effect always increases with the absolute value of the phenotypic optimum. But it does not change if the phenotypic optimum shifts between z_o and $-z_o$. The threshold effect is also larger when selection is weaker ($Ns\bar{\gamma}^2 \ll 2$) or the quantitative trait is controlled mostly by small-effect loci. For phenotypic optimum at zero, (A7.4) gives

$$\hat{\gamma}_N(L) \approx \hat{\gamma}_N + \frac{\hat{\gamma}_N}{4Ns\kappa_2} \tag{A7.6}$$

so that the deviation $\hat{\gamma}_N(L) - \hat{\gamma}_N$ is of order L^{-1}.

For small z_o and large L, (A7.5) shows that the threshold effect does not differ much from the infinite-loci result (21) when selection is strong ($Ns\bar{\gamma}^2 \gg 2$). Furthermore, (A7.5) predicts that the threshold effect for large but finite number of loci is always larger than $\hat{\gamma}_N$, and increases with the magnitude of the phenotypic optimum. Thus, a locus classified as a large-effect locus for a phenotypic optimum at zero can become a small-effect locus if the phenotypic optimum is large enough. This is because for large, positive (negative) phenotypic optimum, population will be well-adapted if the $+$ ($-$) allele's frequency at most loci is close to fixation (that is, the distribution is unimodal and heavily skewed towards high frequency).

CRediT authorship contribution statement

Archana Devi: Conceptualization, Formal analysis, Investigation, Methodology, Software, Visualization, Writing - review and editing. **Kavita Jain:** Conceptualization, Formal analysis, Investigation, Methodology, Supervision, Writing - original draft, Writing - review and editing.

References

, .

Barghi, N., Hermisson, J., Schlötterer, C., 2020. Polygenic adaptation: a unifying framework to understand positive selection. Nat Rev Genet 21, 769–781.

Barton, N.H., 1986. The maintenance of polygenic variation through a balance between mutation and stabilizing selection. Genet. Res. 47, 209–216.

Barton, N.H., Rouhani, S., 1987. The frequency of shifts between alternative equilibria. J. theor. Biol. 125, 397–418.

Bell, G., 2009. The oligogenic view of adaptation. Cold Spring Harbor Symposia on Quantitative Biology 74, 139–144. doi:10.1101/sqb.2009.74.004.

Blinnikov, S., Moessner, R., 1998. Expansions for nearly Gaussian distributions. Astron. Astrophys. Suppl. Ser. 130, 193–205.

Boyle, E.A., Li, Y.I., Pritchard, J.K., 2017. An expanded view of complex traits: from polygenic to omnigenic. Cell 169, 1177–1186.

Bulmer, M.G., 1971. The effect of selection on genetic variability. The American Naturalist 105, 201–211. URL: https://doi.org/10.1086/282718, doi:10.1086/282718.

Bulmer, M.G., 1972. The genetic variability of polygenic characters under optimizing selection, mutation and drift. Genet. Res., Camb. 19, 17–25.

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

Bürger, R., 2000. The Mathematical Theory of Selection, Recombination, and Mutation. Wiley, Chichester.

Courau, P., Lambert, A., Schertzer, E., 2026. Stabilizing selection on a polygenic trait from the gene's-eye view. bioRxiv , 2026.02.23.706325URL: https://www.biorxiv.org/content/10.64898/2026.02.23.706325v1, doi:10.1101/2026.02.23.706325.

García-González, J., O'Reilly, P.F., 2026. The polygenic, omnigenic and stratagenic models of complex disease risk. Nature Genetics 58, 253–263. URL: https://www.nature.com/articles/s41588-025-02467-w, doi:10.1038/s41588-025-02467-w.

Gardiner, C.W., 1997. Handbook of stochastic methods for physics, chemistry and the natural sciences. Springer-Verlag.

Goddard, M.E., Hayes, B.J., 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat. Rev. Genet. 10, 381–391.

Götsch, H., Bürger, R., 2024. Polygenic dynamics underlying the response of quantitative traits to directional selection. Theoretical Population Biology 158, 21–59. doi:10.1016/j.tpb.2024.04.006.

Hahl, S.K., Kremling, A., 2016. A comparison of deterministic and stochastic modeling approaches for biochemical reaction systems: On fixed points, means, and modes. Front. Genet. 7, 1–11.

Hayward, L.K., Sella, G., 2022. Polygenic adaptation after a sudden change in environment. eLife 11, e66697.

Higham, D.J., 2001. An algorithmic introduction to numerical simulation of stochastic differential equations. SIAM Rev. 43, 525–546.

Hill, W., Goddard, M., Visscher, P., 2008. Data and theory point to mainly additive genetic variance for complex traits. PLoS Genet 4, e1000008.

Höllinger, I., Wölfl, B., Hermisson, J., 2023. A theory of oligogenic adaptation of a quantitative trait. Genetics 225, iyad139. URL: https://doi.org/10.1093/genetics/iyad139, doi:10.1093/genetics/iyad139.

Hutzenthaler, M., Jentzen, A., Kloeden, P.E., 2011. Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 467, 1563–1576. URL: https://doi.org/10.1098/rspa.2010.0348, doi:10.1098/rspa.2010.0348.

Jain, K., Stephan, W., 2015. Response of polygenic traits under stabilising selection and mutation when loci have unequal effects. G3: Genes, Genomes, Genetics 5, 1065–1074.

Jain, K., Stephan, W., 2017a. Modes of rapid polygenic adaptation. Mol. Biol. Evol 34, 3169–3175.

Jain, K., Stephan, W., 2017b. Rapid adaptation of a polygenic trait after a sudden environmental shift. Genetics 206, 389–406.

van Kampen, N.G., 1997. Stochastic processes in physics and chemistry. North Holland Personal Library.

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

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gibert, P., Beerli, P., 2001. The strength of phenotypic selection in natural populations. Am. Nat. 157, 245–261.

Lande, R., 1976. Natural selection and random genetic drift in phenotypic evolution. Evolution 30, 314–334.

Latter, B.D.H., 1960. Natural selection for an intermediate optimum. Australian Journal of Biological Sciences 13, 30–35.

Mackay, T.F.C., 2004. The genetic architecture of quantitative traits: lessons from *drosophila*. Current Opinion in Genetics and Development 14, 253–257.

Olver, F.W.J., Olde Daalhuis, A.B., Lozier, D.W., Schneider, B.I., Boisvert, R.F., Clark, C.W., Miller, B.R., Saunders, B.V., Cohl, H.S., McClain, M.A., 2024. NIST Digital Library of Mathematical Functions. http://dlmf.nist.gov/, Release 1.2.3 of 2024-12-15.

Risken, H., 1996. The Fokker Planck equation. Methods of solution and applications. Springer, Berlin.

Robertson, A., 1956. The effect of selection against extreme deviants based on deviation or on homozygosis. Journal of Genetics 54, 236–248.

Rockman, M.V., 2012. The QTN program and the alleles that matter for evolution: all that's gold does not glitter. Evolution 66-1, 1–17.

Sanjaka, J.S., Sidorenkoc, J., Robinson, M.R., Thornton, K.R., Visscher, P.M., 2018. Evidence of directional and stabilizing selection in contemporary humans. Proc. Natl. Acad. Sci. USA 115, 151–156.

Sella, G., Barton, N., 2019. Thinking about the evolution of complex traits in the era of Genome-Wide Association Studies. Annu Rev Genomics Hum Genet 20, 461–493.

Sodeland, M., Jentoft, S., Jorde, P.E., Mattingsdal, M., Albretsen, J., Kleiven, A.R., Wårøy Synnes, A.E., Espeland, S.H., Olsen, E.M., Andrè, C., Stenseth, N.C., Knutsen, H., 2022. Stabilizing selection on Atlantic cod supergenes through a millennium of extensive exploitation. Proc. Natl. Acad. Sci 119, e2114904119.

Thornton, K., 2019. Polygenic adaptation to an environmental shift: temporal dynamics of variation under Gaussian stabilizing selection and additive effects on a single trait. Genetics 213, 1513–1530.

Timpson, N.J., Greenwood, C.M.T., Soranzo, N., Lawson, D.J., Richards, J.B., 2018. Genetic architecture: the shape of the genetic contribution to human traits and disease. Nat. Rev. Genet. 19, 110–124.

Turelli, M., 1988. Phenotypic evolution, constant covariances, and the maintenance of additive variance. Evolution 42, 13421347.

Turelli, M., Barton, N., 1990. Dynamics of polygenic characters under selection. Theor. Popul. Biol. 38, 1–57.

Turelli, M., Barton, N., 1994. Genetic and statistical analyses of strong selection on polygenic traits: what, me normal? Genetics 138, 913941.

de Villemereuil, P., Charmantier, A., Arlt, D., Bize, P., Brekke, P., Brouwer, L., Cockburn, A., Côté, S.D., Dobson, F.S., Evans, S.R., Festa-Bianchet, M., Gamelon, M., Hamel, S., Hegelbach, J., Jerstad, K., Kempenaers, B., Kruuk, L.E.B., Kumpula, J., Kvalnes, T., McAdam, A.G., McFarlane, S.E., Morrissey, M.B., Pärt, T., Pemberton, J.M., Qvarnström, A., Røstad, O.W., Schroeder, J., Senar, J.C., Sheldon, B.C., van de Pol, M., Visser, M.E., Wheelwright, N.T., Tufto, J., Chevin, L.M., 2020. Fluctuating optimum and temporally variable selection on breeding date in birds and mammals. Proc. Natl. Acad. Sci. U.S.A. 117, 31969–31978.

Visscher, P.M., Wray, N.R., Zhang, Q., Sklar, P., McCarthy, M.I., Brown, M.A., Yang, J., 2017. 10 years of GWAS discovery: Biology, function, and translation. Am J Hum Genet. 101, 5–22.

de Vladar, H.P., Barton, N.H., 2014. Stability and response of polygenic traits to stabilizing selection and mutation. Genetics 197, 749–767.

Walsh, B., Lynch, M., 2018. Evolution and Selection of Quantitative Traits. Oxford University Press, Oxford, UK. First edition.

Wright, S., 1937. The distribution of gene frequencies in populations. Proc. Natl. Acad. Sci USA 23, 307–320.