

1 A White Noise Approach to Evolutionary Ecology

2 Bob Week, Steve Krone, Scott L. Nuismer, Luke J. Harmon

3 3/1/2020

4 **Abstract**

5 We derive the dynamics of the distribution of a quantitative character and the abundance of a bio-
6 logical population from a stochastic partial differential equation driven by space-time white noise. In the
7 process we develop a useful set of heuristics to operationalize the powerful, but abstract theory of white
8 noise and measure-valued Markov processes. This approach allows us to compute the full implications
9 of a stochastic process such as demographic stochasticity on phenotypic distributions and abundances
10 of populations. We demonstrate the utility of our approach by deriving a model of diffuse coevolution
11 mediated by exploitative competition for a continuum of resources. Other than trait and abundance dis-
12 tributions, this model predicts interaction networks parameterized by rates of interactions, competition
13 coefficients, and selection gradients. We briefly investigate the relationship between selection gradients
14 and competition coefficients. This illustrative investigation suggests selection gradients can be either
15 positively or negatively correlated with competition coefficients depending on the ratio of interspecific
16 trait variation to intraspecific trait variation. Hence, this approach can contribute to the development of
17 a synthetic theory of evolutionary ecology by formalizing first principle derivations of dynamical equa-
18 tions describing populations and communities which can then be used for rigorous investigations of the
19 relationship between feedbacks of ecological and evolutionary processes and the patterns of diversity they
20 produce.

21 **1 Introduction**

22 Our goal in this manuscript is to develop a rigorous, but accessible approach to synthesize the stochas-
23 tic dynamics of abundance, mean trait and heritable variation in biological populations for the study of
24 theoretical evolutionary ecology. A primary aim of theoretical evolutionary ecology is the development of
25 mathematical approaches to describe the evolution of populations and their interactions with both the biotic
26 and abiotic environments in which they are embedded. Given this consideration, a natural scope for such
27 an approach centers on quantifying the abundance dynamics of populations and the evolution of traits me-
28 diating their interactions as functions of relevant abiotic factors. Although taking into account abundance,
29 phenotype and environment provides the basis for a partial understanding of the complex nature of biolog-
30 ical communities, a deeper understanding must account for the effects of contemporary dispersal and the
31 phylogeographic history of interacting lineages (Kraft et al. 2007; Hickerson et al. 2010; Manceau, Lambert,
32 and Morlon 2016; McPeek 2017) along with the genetic basis of ecologically relevant traits (Conner 2004;
33 Fussman, Loreau, and Abrams 2007) and feedbacks between populations and the biogeochemical cycles they
34 ultimately depend on (Loreau 2010; Ågren and Andersson 2012). It is therefore ideal that the development
35 of any such mathematical approach anticipates extensions to account for these important factors shaping
36 ecological communities, especially as empirical and conceptual work in these directions continues to grow
37 (Abdala-Roberts and Mooney 2014; Kölzsch et al. 2015; Crutsinger 2015; Fitzpatrick et al. 2015, 2017;
38 Marx et al. 2017; Rudman et al. 2017; Skovmand et al. 2018; Nuland et al. 2019; Harmon et al. 2019). Fur-
39 thermore, the approach would benefit from a stochastic component to capture the chance nature of biological
40 reality (Lande, Engen, and SÆther 2003; Meester et al. 2018; Mubayi et al. 2019) and serve as a basis for
41 the construction of statistical methods that measure evolutionary and ecological processes occurring in the
42 wild. Such methods will tether theory to reality and allow for rigorous tests of hypotheses on the structure

43 and behavior of ecological communities. In this paper we introduce a framework that establishes a formal
44 connection between the continuous-time dynamics of abundance and quantitative traits in stochastically
45 evolving populations. We then demonstrate the utility of our framework through the derivation and analysis
46 of a model of diffuse coevolution and discuss how it can be extended to account for the details mentioned
47 above.

48 Current theoretical approaches to synthesize evolution and ecology have capitalized on the fact that biological
49 fitness plays a key role in determining both sets of dynamics. While correlation of fitness and genotype is
50 the basis of evolution by natural selection, the mean fitness across all individuals in a population determines
51 the growth, stasis or decline of abundance. In section 2.1 we review the mathematical formalization of
52 this connection, which has been established in the contexts of population genetics (Crow and Kimura 1970;
53 Roughgarden 1979), evolutionary game theory (Hofbauer and Sigmund 1998; Nowak 2006; Lion 2018),
54 quantitative genetics (Lande 1982; Doebeli 1996; Lion 2018) and a unifying framework for these three
55 distinct approaches to evolutionary theory (Champagnat, Ferrière, and Méléard 2006) which is intimately
56 related to the approach we take here.

57 Reviewing these accomplishments reveals a beautiful synthesis of evolution and population ecology. However,
58 it also reveals a gap in theoretical approaches to incorporate the intrinsically random nature of populations.
59 Specifically, in theoretical quantitative genetics the derivation of a population’s response to random genetic
60 drift is derived in discrete time under the assumption of constant effective population size using arguments
61 based on properties of random samples (Lande 1976). Though this approach conveniently mimics the formalism
62 provided by the Wright-Fisher model of population genetics, real population sizes fluctuate over time.
63 Furthermore, since these fluctuations are themselves stochastic, it seems natural to derive expressions for the
64 evolutionary response to demographic stochasticity and consider how the results relate to characterizations
65 of random genetic drift. This has been done in continuous time for population genetic models without too
66 much technical overhead, assuming a finite number of alleles. However, for populations with a continuum of
67 types, such as a quantitative trait, this becomes a vexing mathematical challenge. Here we close this gap by
68 combining the calculus of white noise with results on rescaled limits of branching Brownian motion processes
69 (BBM) and stochastic partial differential equations (SPDE). Our goal has two components: 1) Establish
70 a novel synthetic approach to theoretical evolutionary ecology that provides a formal connection between
71 demographic stochasticity and random genetic drift in the context of quantitative traits. To show that our
72 approach can be used to develop useful biological insights we derive a model of coevolution in an ecological
73 network and use it to investigate the relationship between competition coefficients and selection gradients.
74 2) Communicate some useful properties of space-time white noise, BBM and SPDE to as wide of audience
75 as possible. With this goal in mind we will not provide a rigorous treatment of any of these deep subjects.
76 Instead, we introduce a set of heuristics that only require the basic concepts of Riemann integration, partial
77 differentiation and some exposure to Brownian motion and stochastic ordinary differential equations (SDE).
78 For a concise introduction to SDE and Brownian motion, we recommend the primer by Evans (2014). Rigorous
79 treatments of SPDE and rescaled limits of BBM can be found in Walsh (1986) and Dawson (1993)
80 respectively.

81 To provide motivation for the stochastic equations developed later and background for our model of co-
82 evolution, we begin with §2.1 by briefly summarizing derivations of deterministic dynamics of biological
83 populations. Starting with a partial differential equation (PDE), we arrive at a general set of ordinary
84 differential equations modelling the dynamics of abundance, trait mean and trait variance. From this we
85 observe that replacing the PDE with a SPDE provides a path to derive SDE describing the evolutionary
86 response to demographic stochasticity. We accomplish this in §2.2 by introducing a set of mathematical tools
87 based on the calculus of white noise and discuss how a diffusion limit of a spatially structured branching
88 process leads to the natural SPDE appropriate for our study. The diffusion limit in turn provides a rigorous
89 method for constructing fitness functions used in models of evolutionary ecology. We employ these tools
90 to derive a system of SDE generalizing our deterministic results to account for demographic stochasticity.
91 However, although biologically insightful, these equations remain difficult to analyze and implement numerically.
92 In §2.3 we use an assumption of normally distributed trait values to simplify these expressions into
93 formulae that are much easier to work with. We then account for the constraint of adaptive evolution on the
94 amount of heritable variation in a population by extending these results via a model of imperfect inheritance.
95 The resulting equations coincide with classical results in quantitative genetics as a special case. In §3 we

96 combine the derived equations of population dynamics with classical niche theory to formulate a model of
 97 coevolution across a guild of S species participating in exploitative competition along a common resource
 98 continuum. In SM §5.7 we apply a classical theorem on rescaled limits of BBM that allow for ecological
 99 interactions to provide a rigorous derivation. To gain biological insight, in §3.2 we numerically integrate our
 100 model of coevolution for $S = 100$ species, tracking the dynamics of traits and abundances, under scenarios
 101 of weak and strong competition. We include an account of the natural history of the simulated system
 102 and discuss the significance of demographic stochasticity for structuring ecological communities. In §3.3
 103 we provide formula for selection gradients and competition coefficients implied by our model and use these
 104 to investigate the relationship between the strengths of competition and coevolution. Finally, §4 concludes
 105 with a summary of accomplishments, a few remarks on the limits of this approach and future directions to
 106 incorporate more explicitly the genetic architecture of populations, feedbacks with ecosystem processes and
 107 the macroevolutionary history of interacting lineages.

108 2 The framework

109 At the core of our approach is a stochastic analog of the replicator equation with mutation in continuous
 110 time and phenotypic space (Taylor and Jonker 1978; Schuster and Sigmund 1983). From this stochastic
 111 replicator-mutator equation we derive a system of SDE for the dynamics of abundance, mean trait and
 112 additive genetic variance of a population. Hence, our approach develops a quantitative genetic theory of
 113 evolutionary ecology. A popular alternative to quantitative genetics is the theory of adaptive dynamics. As
 114 demonstrated by Page and Nowak (2002), the canonical equation of adaptive dynamics can be derived from
 115 the replicator-mutator equation. Thus, one could start from the atomic roots of our approach and pursue
 116 a stochastic adaptive dynamic theory instead. We choose the former in anticipation of an extension of our
 117 approach that explicitly models the genetic details of populations.
 118 In this section we review the derivations of the replicator-mutator equation and trait dynamics from abun-
 119 dance dynamics and extend these formulae along with related results to the stochastic case. The results
 120 established in this section provide the framework from which larger scale ecological stuctures, such as species
 121 abundance distributions and interaction networks, can be computed.

122 2.1 Deterministic dynamics

123 Our review begins by considering the dynamics of an asexually reproducing population in a homogeneous
 124 environment. Individuals are assumed to be haploid and carry one of K alleles each with a different fitness
 125 expressed as growth rate. Under these assumptions, the derivation of the evolution of allele frequencies
 126 due to natural selection can be derived from expressions of exponential growth. This, and a few related
 127 approaches, have been provided by Crow and Kimura (1970, §5.3). Specifically, denoting ν_i the abundance
 128 of individuals with allele i and m_i the growth rate of allele i (called the Malthusian parameter in Crow and
 129 Kimura 1970), we have

$$\frac{d\nu_i}{dt} = m_i \nu_i. \quad (1)$$

130 Starting from this model, we get the total abundance of the population as $N = \sum_{i=1}^K \nu_i$, the frequency of
 131 allele i as $p_i = \nu_i/N$ and the mean fitness of the population as $\bar{m} = \sum_{i=1}^K p_i m_i$. Hence, we can employ
 132 some elementary calculus to derive the dynamics of abundance dN/dt and the dynamics of allele frequencies
 133 $dp_1/dt, \dots, dp_K/dt$ as

$$\frac{dN}{dt} = \sum_{i=1}^K \nu_i m_i = N \sum_{i=1}^K p_i m_i = \bar{m} N, \quad (2)$$

$$\frac{dp_i}{dt} = \frac{d}{dt} \frac{\nu_i}{N} = \frac{1}{N^2} \left(N \frac{d\nu_i}{dt} - \frac{dN}{dt} \nu_i \right) = \frac{1}{N} (m_i \nu_i - \bar{m} N p_i) = (m_i - \bar{m}) p_i. \quad (3)$$

135 Two important observations of these equations include 1) mean fitness \bar{m} determines the abundance dynamics
 136 of the entire population and 2) allele i will increase (decrease) in frequency if $m_i > \bar{m}$ ($< \bar{m}$). Equation (3)
 137 is known in the field of evolutionary game theory as the replicator equation (Hofbauer and Sigmund 1998;
 138 Nowak 2006; Lion 2018; Taylor and Jonker 1978; Schuster and Sigmund 1983). Instead of being explicitly
 139 focused on alleles, the replicator equation describes the fluctuations of relative abundances of various *types*
 140 in a population in terms of the vital rates of each type. Using a matrix of transition rates between differing
 141 types, it is straight-forward to extend the replicator equation to include mutation, which is known as the
 142 replicator-mutator equation (Nowak 2006).

143 Inspired by equations (1)-(3), we derive an analog of the replicator-mutator equation for a continuum of
 144 types (that is, for a quantitative trait). In particular, we model a continuously reproducing population
 145 with trait values $x \in \mathbb{R}$ and an abundance density $\nu(x, t)$ that represents the amount of individuals in the
 146 population with trait value x at time t . Hence, the abundance density satisfies $N(t) = \int_{-\infty}^{+\infty} \nu(x, t) dx$ and
 147 $p(x, t) = \nu(x, t)/N(t)$ is the relative density of trait x which we also refer to as the phenotypic distribution.

148 To stay within the realm of biological plausibility we require a set technical assumptions. First, we assume
 149 the initial abundance density is continuous, non-negative, integrable and has finite trait mean and variance.
 150 That is, we assume $\nu(x, 0)$ is continuous in x , satisfies $\nu(x, 0) \geq 0$ for all $x \in \mathbb{R}$ and

$$N(0) = \int_{-\infty}^{+\infty} \nu(x, 0) dx < +\infty, \quad (4)$$

$$-\infty < \bar{x}(0) = \int_{-\infty}^{+\infty} xp(x, 0) dx < +\infty, \quad (5)$$

$$\sigma^2(0) = \int_{-\infty}^{+\infty} (x - \bar{x}(0))^2 p(x, 0) dx < +\infty, \quad (6)$$

151 where $\bar{x}(t)$ and $\sigma^2(t)$ are respectively the mean trait and phenotypic variance at time $t \geq 0$. We denote
 152 by $C_{1,c}^+(\mathbb{R} \times [0, \infty))$ the set of all non-negative integrable functions that are continuous in x and t . Second,
 153 we assume selection is determined by the growth rate $m(\nu, x)$ that is differentiable with respect to $x \in \mathbb{R}$,
 154 continuous with respect to $\nu \in C_{1,c}^+(\mathbb{R} \times [0, \infty))$ and satisfies $m(\nu, x) \leq r$ for some $r \in \mathbb{R}$ and all $x \in \mathbb{R}$ and
 155 $\nu \in C_{1,c}^+(\mathbb{R} \times [0, \infty))$. Our continuity assumption implies; if a sequence of abundance densities $\{\nu_n\}_{n=1}^\infty \subset$
 156 $C_{1,c}^+(\mathbb{R} \times [0, \infty))$ satisfies $\lim_{n \rightarrow \infty} \nu_n = \nu \in C_{1,c}^+(\mathbb{R} \times [0, \infty))$, then $\lim_{n \rightarrow \infty} m(\nu_n, x) = m(\nu, x)$ for all $x \in \mathbb{R}$.
 157 Third, we assume mutation is captured by diffusion with coefficient $\frac{\mu}{2}$. With these technicalities aside, the
 158 demographic dynamics can be modelled by the PDE

$$\frac{\partial}{\partial t} \nu(x, t) = m(\nu, x) \nu(x, t) + \frac{\mu}{2} \frac{\partial^2}{\partial x^2} \nu(x, t) \quad (7)$$

159 with the initial condition $\nu(x, 0)$ described above. This PDE is semilinear due to the dependency of the
 160 growth rate $m(\nu, x)$ on the solution $\nu(x, t)$ and is referred to as a scalar reaction-diffusion equation (Evans
 161 2010). When $\mu = 0$, equation (7) can be seen as an analog of equation (1) for a continuum of types. By
 162 assuming mutation acts as diffusion the effect of mutation causes $\nu(x, t)$ to flatten out over time. In fact,
 163 if the growth rate is constant across x , then this model of mutation will cause $\nu(x, t)$ to converge to a flat
 164 line as $t \rightarrow \infty$. Although clearly an idealized representation of biological reality, this model is sufficiently
 165 general to capture a large class of dynamics including density dependent growth and frequency dependent
 166 selection. As an example, logistic growth combined with quadratic stabilizing selection can be captured
 167 using the growth rate $m(\nu, x) = r - \frac{a}{2}(\theta - x)^2 - c\nu(x, t)$ where $a \geq 0$ is the strength of stabilizing selection
 168 around the phenotypic optimum $\theta \in \mathbb{R}$, $c \geq 0$ captures the effect of intraspecific competition and $r \in \mathbb{R}$ is
 169 the intrinsic growth rate in the absence of abiotic selection.

170 To derive a replicator-mutator equation from equation (7), we employ the chain rule from calculus. Writing
 171 $\bar{m}(t) = \int_{-\infty}^{+\infty} m(\nu, x)p(x, t)dx$ for the mean fitness, we have

$$\begin{aligned}
\frac{d}{dt}N(t) &= \frac{d}{dt}\int_{-\infty}^{+\infty}\nu(x,t)dx = \int_{-\infty}^{+\infty}\frac{\partial}{\partial t}\nu(x,t)dx \\
&= \int_{-\infty}^{+\infty}m(\nu,x)\nu(x,t)dx + \int_{-\infty}^{+\infty}\frac{\mu}{2}\frac{\partial^2}{\partial x^2}\nu(x,t)dx \\
&= N(t)\int_{-\infty}^{+\infty}m(\nu,x)p(x,t)dx = \bar{m}(t)N(t). \quad (8)
\end{aligned}$$

174 By our assumptions on mutation and rate of growth, $\nu(x,t)$ is twice differentiable with respect to x and
175 $\int_{-\infty}^{+\infty}\nu(x,t)dx < \infty$ for all $t \geq 0$ (SM §§??). This implies that we are justified in swapping the order of
176 differentiation and integration and the result $\int_{-\infty}^{+\infty}\frac{\partial^2}{\partial x^2}\nu(x,t)dx = 0$ can be derived from the fundamental
177 theorem of calculus. Biological reasoning agrees with this latter result since mutation neither creates nor
178 destroys individuals, but merely changes their type from their parental type. Taking the same approach, we
179 derive the dynamics of the phenotypic distribution $p(x,t)$ in response to selection and mutation as

$$\begin{aligned}
\frac{\partial}{\partial t}p(x,t) &= \frac{\partial}{\partial t}\frac{\nu(x,t)}{N(t)} = \frac{1}{N^2(t)}\left(N(t)\frac{\partial}{\partial t}\nu(x,t) - \nu(x,t)\frac{d}{dt}N(t)\right) \\
&= \frac{1}{N(t)}\left(m(\nu,x)\nu(x,t) + \frac{\mu}{2}\frac{\partial^2}{\partial x^2}\nu(x,t) - \bar{m}(t)\nu(x,t)\right) \\
&= (m(\nu,x) - \bar{m}(t))p(x,t) + \frac{\mu}{2}\frac{\partial^2}{\partial x^2}p(x,t). \quad (9)
\end{aligned}$$

180 This result closely resembles Kimura's continuum-of-alleles model (Kimura 1965; Bürger 1986). The primary
181 difference being that our model utilizes diffusion instead of convolution with an arbitrary mutation kernel.
182 Of course, our model of mutation can be derived as an approximation to Kimura's model, which has been
183 referred to as the Gaussian allelic approximation (in reference to the distribution of mutational effects at
184 a given locus of a genome on the values of traits, Bürger 2000), the infinitesimal model (in reference to
185 modelling continuous traits as being encoded by an infinite number of loci each having infinitesimal effect,
186 Barton, Etheridge, and Véber 2017) and the Gaussian descendants approximation (in reference to offspring
187 trait values being normally distributed around their parental values, Turelli 2017). Alternatively, since
188 diffusion is the continuous-time equivalent to convolution against a Gaussian kernel (SM §5.2), equation (9)
189 can also be seen as a special case of the continuum-of-alleles model.

190 The covariance of fitness and phenotype across the population is defined as

$$\text{Cov}_t(m(\nu,x),x) = \int_{-\infty}^{+\infty}(m(\nu,x) - \bar{m}(t))(x - \bar{x}(t))p(x,t)dx. \quad (10)$$

191 Hence, the dynamics of the mean trait $\bar{x}(t)$ can be derived as

$$\begin{aligned}
\frac{d}{dt}\bar{x}(t) &= \frac{d}{dt}\int_{-\infty}^{+\infty}xp(x,t)dx = \int_{-\infty}^{+\infty}x\frac{\partial}{\partial t}p(x,t)dx \\
&= \int_{-\infty}^{+\infty}x(m(\nu,x) - \bar{m}(t))p(x,t) + x\frac{\mu}{2}\frac{\partial^2}{\partial x^2}p(x,t)dx \\
&= \text{Cov}_t(m(\nu,x),x) + \frac{\mu}{2}\int_{-\infty}^{+\infty}x\frac{\partial^2}{\partial x^2}p(x,t)dx. \quad (11)
\end{aligned}$$

192 Equation (11) is a continuous time analog of the well known Robertson-Price equation (Robertson 1966;
193 Price 1970; Frank 2012; Queller 2017; Lion 2018). The covariance of fitness and phenotype creates change

194 in \bar{x} to maximize mean fitness \bar{m} . Since this change is driven by a covariance with respect to phenotypic
 195 diversity, the response in mean trait to selection is mediated by the phenotypic variance. In particular,
 196 when $\sigma^2 = 0$, \bar{x} will not respond to selection. The second term, due to mutation, represents the so-called
 197 *transmission bias* (Frank 2012). When $p(x, t)$ is symmetric around \bar{x} , this term disappears due to equal bias
 198 in opposite directions cancelling the effect of mutation on \bar{x} . Following the approach taken to calculate the
 199 evolution of \bar{x} , we find the response of phenotypic variation to this model of selection and mutation is

$$\frac{d}{dt}\sigma^2(t) = \text{Cov}_t(m(\nu, x), (x - \bar{x})^2) + \frac{\mu}{2} \int_{-\infty}^{+\infty} (x - \bar{x})^2 \frac{\partial^2}{\partial x^2} p(x, t) dx. \quad (12)$$

200 For the sake of space we relegate the derivation of $d\sigma^2/dt$ to SM §5.3. In the absence of mutation equation
 201 (12) agrees with the result derived by Lion (2018) for discrete phenotypes. From a statistical perspective,
 202 if we think of $(x - \bar{x})^2$ as a square error, then in analogy to the dynamics of the mean trait, we see that
 203 the response in σ^2 to selection can be expressed as a covariance of fitness and square error, which is defined
 204 in analogy to $\text{Cov}_t(m(\nu, x), x)$. This covariance also creates change in σ^2 that maximizes mean fitness \bar{m} ,
 205 which can be positive or negative depending on whether selection is stabilizing or disruptive. In SM §5.3 we
 206 show when $p(x, t)$ is a Gaussian bell curve, the increase in variance due to mutation simplifies to μ .

207 In SM §5.5, we extend these results to include the effects of demographic stochasticity. Similar to the
 208 approach taken by Champagnat, Ferrière and Méléard (2006), we begin with a BBM that models populations
 209 as discrete sets of reproducing individuals whose vital rates depend on their trait value as well as the state of
 210 the entire population. Taking a large population size limit and keeping our assumption of single dimensional
 211 traits, we employ a pair of classical results that show, under the appropriate rescaling in time, phenotypic
 212 space and population density, a sequence of rescaled BBM converges to a limiting process that can be
 213 characterized by a SPDE (Méléard and Roelly 1993; Li 1998). The limiting processes of rescaled BBM
 214 have been referred to as measure-valued Markov processes (Dawson 1993) or superprocesses (Etheridge
 215 2000). Under the simplifying assumptions inherited from our treatment of deterministic dynamics and an
 216 additional assumption on demographic stochasticity, we obtain as a special case a relatively simple expression
 217 for an SPDE that generalizes equation (7). The simplicity of our special case allows us to use properties
 218 of space-time white noise processes to derive a set of SDE that generalize equations (8), (11) and (12) to
 219 include the effects of demographic stochasticity. Classical expressions for the effects of random genetic drift
 220 on the evolution of mean traits are obtained as a special case.

221 In the following section we provide the necessary mathematical tools needed to derive SDE from SPDE.
 222 Since our aim is to present this material to as wide of audience as possible, our approach deviates from
 223 standard definitions to remove the need for a detailed technical treatment. In addition to the notions of
 224 Riemann integration and partial differentiation already employed, the reader will only need some elementary
 225 probability and an intuitive understanding of SDE, including Brownian motion. Because space-time white
 226 noise, denoted by $\dot{W}(x, t)$, appears in the SPDE characterizing diffusion limits of BBM, we begin by defining
 227 $\dot{W}(x, t)$ and illustrating its relevant properties including a set of heuristics for performing calculations.
 228 Treating only the simplest of cases, we then provide a brief review of BBM, their diffusion limits and the
 229 SPDE that characterize them. For those not interested in the white noise calculus or superprocesses and
 230 would rather jump straight into more biologically relevant results, we recommend skipping to §2.3.1.

231 2.2 White noise calculus and superprocesses

232 2.2.1 Definition and basic properties of white noise

233 One can think of white noise as the static seen on old television sets or infinitely detailed random dust
 234 spread across both time and space. From a more mathematical, yet still informal perspective, white noise
 235 can be thought of as a stochastic process. That is, we can picture white noise as a collection of random
 236 variables indexed by time and possibly space. In relation to Brownian motion, denoted by W , white noise
 237 can be interpreted of as the derivative of Brownian motion with respect to time, denoted \dot{W} . Since Brownian
 238 motion can be thought to take infinitesimally small Gaussian distributed jumps at each time point, this leads

239 to the conceptualization of white noise as a collection of Gaussian distributed random variables. Figure 1
240 illustrates realizations of this conceptualized white noise in one (left) and two (right) dimensions.

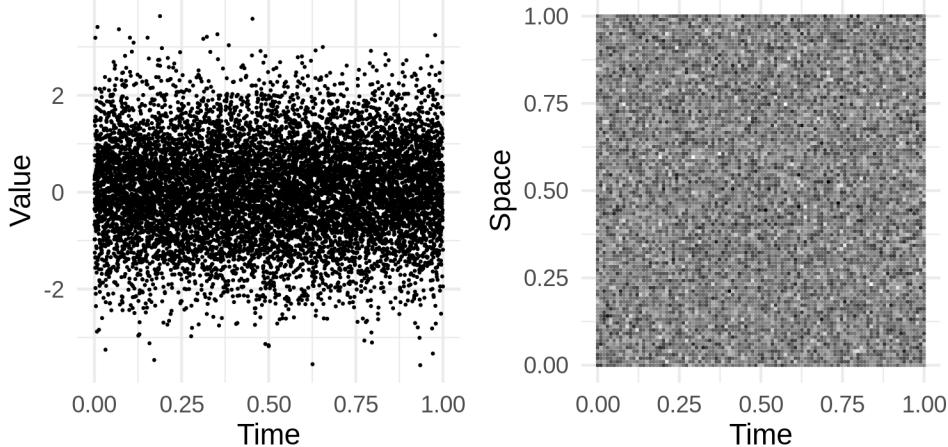


Figure 1: Approximations of sample paths of temporal white noise (left) and space-time white noise (right) with brightness scaled to value.

241 However, it turns out that realizations of white noise do not exist as functions in the classical sense. Indeed,
242 since Brownian motion is nowhere differentiable with respect to time, white noise cannot be formally under-
243 stood as its time derivative. Thus our notation \dot{W} is only meant to aid intuition and not be taken as formal.
244 A formal understanding is possible by considering white noise as a *measure*-valued process (Dawson 1975;
245 Walsh 1986) or as a *generalized* process that acts on classically defined functions or stochastic processes to
246 return random variables or stochastic processes (Krylov and Rozovskii 1981; Da Prato and Zabczyk 2014).
247 Since a measure-valued process can be defined from a generalized process and a generalized process can be
248 defined from a measure-valued process, the distinction between the two is more or less a matter of perspec-
249 tive. However, we find the perspective of white-noise as a generalized process to be a more efficient route for
250 developing heuristics to help with some routine calculations. Hence, this latter approach provides the general
251 idea implemented here. Although the treatments provided by Krylov and Rozovskii (1981) and Da Prato
252 and Zabczyk (2014) extend the theory of SDE to formally treat SPDE in a general and elegant fashion, they
253 require the navigation of an enormous amount of technical definitions and detailed proofs. To extract some
254 particularly useful results from this theory relevant to our goal of synthesizing the stochastic dynamics of
255 biological populations, we provide a streamlined approach by capitolizing on the solid ground these authors
256 have established. For instance, instead of rigorously proving properties of white-noise, we simply define them
257 to be so, taking solice in the fact that the technical details have been worked out elsewhere.

258 To shed a bit of light on the idea of a generalized process, we note that they are the stochastic analog of a
259 generalized function, such as the Dirac delta function δ . Just as a generalized function operates on classical
260 functions to return a value (e.g., $\delta(f) = f(0)$), a generalized process acts on a set of functions to return
261 a classically defined stochastic process. For a breif overview of the theory of generalized functions, see the
262 addendum to chapter 3 of Kolmogorov and Fomin (1999).

263 Throughout this section, we minimize notation by writing $\int_{\mathbb{R}} f(x)dx = \int_{-\infty}^{+\infty} f(x)dx$ and similarly $\int_D f(x)dx$
264 for the integral of f over $D \subset \mathbb{R}$. We define \mathcal{N}_2 as the set of stochastic processes $f(x, t)$ that are continuous
265 in both x and t and satisfy $\mathbb{E} \left(\int_0^t \int_{\mathbb{R}} f^2(x, s) dx ds \right) < +\infty$ for each $t \geq 0$. The operator \mathbb{E} denotes expectation
266 with respect to the underlying probability space. For each $t \geq 0$ we set

$$\|f\|_t = \sqrt{\mathbb{E} \left(\int_0^t \int_{\mathbb{R}} f^2(x, s) dx ds \right)}. \quad (13)$$

267 We define a generalized stochastic process \mathbf{W} that maps processes $f \in \mathcal{N}_2$ to real-valued stochastic processes
 268 indexed by time $t \geq 0$, but not by space. To evaluate \mathbf{W} for a process $f \in \mathcal{N}_2$ and some time $t \geq 0$ we write
 269 $\mathbf{W}_t(f)$. Specifically, for any $f, g \in \mathcal{N}_2$, we define $\mathbf{W}(f)$ and $\mathbf{W}(g)$ to be Gaussian processes satisfying, for
 270 any $t, t_1, t_2 \geq 0$,

$$\mathbb{E}(\mathbf{W}_t(f)) = \mathbb{E}(\mathbf{W}_t(g)) = 0, \quad (14)$$

$$\mathbb{C}(\mathbf{W}_{t_1}(f), \mathbf{W}_{t_2}(g)) = \mathbb{E}\left(\int_0^{t_1 \wedge t_2} \int_{\mathbb{R}} f(x, s) g(x, s) dx ds\right), \quad (15)$$

272 where $t_1 \wedge t_2 = \min(t_1, t_2)$ and \mathbb{C} denotes covariance with respect to the underlying probability space. The
 273 operators \mathbb{E} and \mathbb{C} are to be distinguished from $\bar{f}(t)$ and $\text{Cov}_t(f, g)$ which denote expectation and covariance
 274 with respect to phenotypic diversity at time $t \geq 0$.

275 Since Gaussian processes are characterized by their expectations and covariances and since we assume the
 276 \mathcal{N}_2 processes are continuous in both time and space, the processes $\mathbf{W}(f)$ and $\mathbf{W}(g)$ are well defined. In
 277 particular, if $f \in \mathcal{N}_2$ is independent of time, then $\mathbf{W}(f)$ is a Brownian motion with variance at time $t \geq 0$
 278 equal to $t \mathbb{E}(\int_{\mathbb{R}} f^2(x, 0) dx)$. With the generalized process \mathbf{W} defined, we define the space-time white noise
 279 $\dot{W}(x, t)$ implicitly via the stochastic integral

$$\left\langle \int_0^t \int_{\mathbb{R}} f(x, s) \dot{W}(x, s) dx ds \right\rangle = \left\langle \int_{\mathbb{R}} \int_0^t f(x, s) \dot{W}(x, s) dx ds \right\rangle = \mathbf{W}_t(f), \quad \forall f \in \mathcal{N}_2, t \geq 0. \quad (16)$$

280 We place quotations in the above expression to emphasize its informal nature and that it should not be
 281 confused with classical Riemann integration. Following this definition of white noise, we compute its value
 282 by sampling it using $L_c^2(\mathbb{R} \times [0, \infty))$ functions. For example, integrating white noise over a region $D \times [0, t]$,
 283 with $t > 0$ and D a bounded subset of \mathbb{R} , is equivalent to evaluating $\mathbf{W}_t(I_{D \times [0, \infty)})$ where

$$I_{D \times [0, \infty)}(x, t) = \begin{cases} 0, & x \notin D \\ 1, & x \in D \end{cases}. \quad (17)$$

284 Since $\int_{\mathbb{R}} I_{D \times [0, \infty)}^2(x, t) dx = \int_D dx = |D|$, where $|D|$ denotes the length of D , $I_{D \times [0, \infty)} \in L_c^2(\mathbb{R} \times [0, \infty))$.
 285 Thus, by equations (14) and (15), we have

$$\mathbb{E}\left(\int_0^t \int_D \dot{W}(x, s) dx ds\right) = 0, \quad (18)$$

$$\mathbb{V}\left(\int_0^t \int_D \dot{W}(x, s) dx ds\right) = t|D|, \quad (19)$$

286 where \mathbb{V} denotes the variance operator with respect to the underlying probability space. Using this notation
 287 equations (14) and (15) can be rewritten as

$$\mathbb{E}\left(\int_0^t \int_{\mathbb{R}} f(x, s) \dot{W}(x, s) dx ds\right) = 0, \quad (20)$$

$$\mathbb{C}\left(\int_0^{t_1} \int_{\mathbb{R}} f(x, s) \dot{W}(x, s) dx ds, \int_0^{t_2} \int_{\mathbb{R}} g(x, s) \dot{W}(x, s) dx ds\right) = \int_0^{t_1 \wedge t_2} \int_{\mathbb{R}} f(x, s) g(x, s) dx ds. \quad (21)$$

290 To relate these formulae to the common notation used for SDE, we write $d\hat{\mathbf{W}}_t(f) = \frac{1}{\|f\|_2(t)} (\int_{\mathbb{R}} f(x, t) \dot{W}(x, t) dx) dt$
 291 so that

$$\int_0^t d\hat{\mathbf{W}}_s(f) = \int_0^t \int_{\mathbb{R}} \frac{f(x, s)}{\sqrt{\int_{\mathbb{R}} f^2(s, y) dy}} \dot{W}(x, s) dx ds. \quad (22)$$

292 This implies $\mathbb{E}(\int_0^t d\hat{\mathbf{W}}_s(f)) = 0$, $\mathbb{C}(\int_0^{t_1} d\hat{\mathbf{W}}_s(f), \int_0^{t_2} d\hat{\mathbf{W}}_s(f)) = t_1 \wedge t_2$ and, as a function of t , $\int_0^t d\hat{\mathbf{W}}_s(f)$
 293 is a standard Brownian motion for any $f \in L_c^2(\mathbb{R} \times [0, \infty))$. Hence, $d\hat{\mathbf{W}}_t(f)$ is analogous to the traditional
 294 shorthand used to denote stochastic differentials. Thus, equation (21) effectively extends Itô's multiplication
 295 table to:

Table 1: An extension of Itô’s multiplication table.

	$d\hat{\mathbf{W}}_t(f)$	$d\hat{\mathbf{W}}_t(g)$	dt
$d\hat{\mathbf{W}}_t(f)$	dt	$\frac{(\int_{\mathbb{R}} f g dx) dt}{\ f\ _2 \ g\ _2}$	0
$d\hat{\mathbf{W}}_t(g)$	$\frac{(\int_{\mathbb{R}} f g dx) dt}{\ f\ _2 \ g\ _2}$	dt	0
dt	0	0	0

296 The extension of Itô’s multiplication table and properties of white noise outlined in this subsection provide
 297 a useful set of tools for working with SPDE. In SM §5.5 we employ these tools to derive SDE that track the
 298 dynamics of abundance, mean trait and phenotypic variance of a population from a particular SPDE. In the
 299 following subsection, we review how this particular SPDE naturally arises as the diffusion limit of a BBM.

300 2.2.2 From branching processes to SPDE

301 In real populations individuals are born and potentially reproduce before they ultimately die. These three
 302 events provide the basic ingredients of a branching process. Mathematical investigations of such processes
 303 have a relatively deep history (Kendall 1966). The most simple branching process, known as the Galton-
 304 Watson process, describes the number of individuals alive at a given time $t \geq 0$ as a non-negative integer
 305 (Kimmel and Axelrod 2015). Feller (1951) introduced a formal method to approximate branching processes
 306 with diffusion processes which are continuous in state (i.e., population size is approximated as a continuous
 307 quantity). Since diffusion processes possess greater analytical tractability than branching processes, Feller’s
 308 method, known as the diffusion limit, has acquired immense popularity particularly in the field of mathematical
 309 population genetics (Ewens 2004). For over the past half of a century a great deal of accomplishments
 310 have been achieved in formalizing the diffusion limits of branching processes that describe populations of
 311 individuals occurring in some continuous space (Watanabe 1968; Dawson 1975; Perkins 1992, 1995; Méléard
 312 and Roelly 1993; Li 1998; Bertoin and Le Gall 2003; Etheridge 2008; Barton and Etheridge 2019). This space
 313 can represent geographic space or, relevant to our context, phenotypic space. In the following subsection, we
 314 describe the BBM process, which is a particularly important branching process with spatial structure. This
 315 process has been very useful in the study of SPDE due to its simplifying assumption that individuals do
 316 not interact. However, this assumption imposes an unfortunate restriction by precluding the modelling of
 317 ecological interactions. We therefore follow our discussion of BBM with a review of a few important results
 318 on spatially structured branching processes that account for interactions.

319 Branching Brownian motion

320 A BBM tracks individuals navigating d -dimensional Euclidean space that reproduce and senesce between
 321 exponentially distributed intervals. Unlike other stochastic processes that take values in \mathbb{R}^d , BBM takes
 322 values in the set of *non-negative finite measures* over \mathbb{R}^d . Intuitively, one can think of a finite measure as a
 323 function that maps subsets of \mathbb{R}^d to real numbers. To be formal, we only consider the Borel subsets of \mathbb{R}^d
 324 corresponding to the Euclidean metric, but understanding this technicality is not crucial to our discussion.
 325 In particular, denoting X_t a BBM, for a subset $D \subset \mathbb{R}^d$, $X_t(D)$ returns the (random) number of individuals
 326 alive within the region D at time $t \geq 0$. The BBM has three main components:

- 327 1) **Branching rate:** In our formulation of BBM we assume Lifetimes of individuals are exponentially
 328 distributed with death rate $\lambda > 0$ and reproduction occurs simultaneously with death. Biologically, this
 329 implies individuals are semelparous An alternative formulation treats birth and death events separately
 330 to model iteroparity. However, under the appropriate rescaling, both approaches converge to the same
 331 diffusion limit. We therefore choose the former approach for the sake of simplicity.

332 2) **Reproductive law:** When a birth event occurs a random (possibly zero) number of offspring are left.
 333 The distribution of offspring left is called the reproductive law or branching mechanism. We denote
 334 the mean and variance in reproductive output by \mathcal{W} and V respectively. The case of $\mathcal{W} = 1$ is referred
 335 to as the critical condition. Under the critical condition the probability that extinction occurs in finite
 336 time is equal to one.

337 3) **Spatial movement:** Each offspring is born at the current location of their parent. Immediately after
 338 birth they move around space according to d -dimensional Brownian motion with diffusion parameter
 339 $\sqrt{\mu}$. In our context we interpret spatial movement as mutation so that the location of an individual at
 340 death represents the value of its phenotype. Then an individual born at location $x \in \mathbb{R}^d$ that lives for
 341 $\tau > 0$ units of time will have a normally distributed trait centered on x with covariance matrix equal to
 342 $\tau\mu$ times the $d \times d$ identity matrix. Hence, offspring inherit normally distributed traits centered on their
 343 parental trait. This fact creates a vital link to the deterministic dynamics reviewed above. Indeed,
 344 in the absence of selection, the deterministic PDE (7) reduces to the $d = 1$ -dimensional Kolmogorov
 345 forward equation for a Brownian motion with diffusion parameter $\sqrt{\mu}$.

346 To obtain a SPDE from a BBM we take a diffusion limit. There are several ways to do this, but a simple
 347 approach is to rescale the mass of individuals and time by $1/n$, diffusion by $\mu \rightarrow \mu/n$, branching rate by
 348 $\lambda \rightarrow n\lambda$, fitness by $\mathcal{W} \rightarrow \mathcal{W}^{1/n}$ and consider the limit as $n \rightarrow \infty$. Denoting the rescaled process by $X_t^{(n)}(D)$,
 349 the limiting process $\mathcal{X}_t = \lim_{n \rightarrow \infty} X_t^{(n)}$ is called a super-Brownian motion and is also a non-negative finite
 350 measure-valued process (Watanabe 1968). However, instead of returning the number of individuals alive in a
 351 region of space, super-Brownian motion returns the *mass* of the population concentrated in a region of space.
 352 Since we have rescaled individual mass by $1/n$ and took the limit $n \rightarrow \infty$, individuals are no longer discrete
 353 units. The particle view of the population gets replaced by a continuous blob spread across d -dimensional
 354 space. In particular, the value of $\mathcal{X}_t(D)$ is a continuously varying non-negative random variable for any
 355 $t \geq 0$ and $D \subset \mathbb{R}^d$. It turns out that for spatial dimension $d = 1$, \mathcal{X}_t is absolutely continuous with respect to
 356 the Lebesgue measure for each $t \geq 0$ (Konno and Shiga 1988; Reimers 1989). This means that we can write
 357 $\mathcal{X}_t(D) = \int_D \nu(x, t) dx$ for some density process $\nu(x, t)$. Setting $\lambda = 1$ and $m = \ln \mathcal{W}$ this density process
 358 satisfies the SPDE

$$\frac{\partial}{\partial t} \nu(x, t) = m\nu(x, t) + \frac{\mu}{2} \frac{\partial^2}{\partial x^2} \nu(x, t) + \sqrt{V\nu(x, t)} \dot{W}(x, t). \quad (23)$$

359 Since $\nu(x, t)$ is not generally differentiable in x or t , the derivatives in expression (23) are taken in the
 360 *weak* sense. That is, to rigorously interpret SPDE (23), we integrate the solution $\nu(x, t)$ against functions
 361 $f \in C_b^2(\mathbb{R})$, where $C_b^2(\mathbb{R})$ is the set of bounded and twice continuously differentiable functions on \mathbb{R} . Hence,
 362 equation (23) is just an abbreviation for

$$\begin{aligned} \int_{\mathbb{R}} \nu(x, t) f(x) dx - \int_{\mathbb{R}} \nu(x, 0) f(x) dx &= \int_0^t \int_{\mathbb{R}} \nu(x, s) \left(m f(x) + \frac{\mu}{2} \frac{\partial^2}{\partial x^2} f(x) \right) ds dx \\ &\quad + \int_0^t \int_{\mathbb{R}} f(x) \sqrt{V\nu(x, s)} \dot{W}(x, s) dx ds, \quad \forall f \in C_b^2(\mathbb{R}). \end{aligned} \quad (24)$$

363 Evans (2010) provides a nice introduction to weak derivatives. For more on the general theory of SPDE
 364 see Walsh (1986). Note that since $\nu(x, t)$ is the density of a finite measure, it is integrable for each $t \geq 0$.
 365 Thus, since for some $M > 0$, $|f(x)| \leq M$ for every $x \in \mathbb{R}$, setting $\varphi(x, t) = f(x) \sqrt{V\nu(x, t)}$ implies $\varphi \in$
 366 $L_c^2(\mathbb{R} \times [0, \infty))$. Hence, the white noise integral on the right-hand side of equation (24) can be understood
 367 using the heuristics introduced above. Evaluating equation (24) in the particular case of $f(x) \equiv 1$ returns
 368 the total mass process, which we refer to as the total abundance $N(t)$.

369 A convergence theorem for the diffusion limit of a generalization of BBM was established by Watanabe
 370 (1968). Dawson (1975) suggested that, for spatial dimension $d = 1$, this diffusion limit should admit a density
 371 process that satisfies a SPDE. Konno and Shiga (1988) and Reimers (1989) independently proved Dawson's
 372 suggestion was indeed correct. The diffusion limit of this more general branching process (in arbitrary
 373 spatial dimension) is referred to as the Dawson-Watanabe superprocess (Etheridge 2000). Conditioning the

374 Dawson-Watanabe superprocess to have constant mass returns the Fleming-Viot process (Etheridge and
 375 March 1991; Perkins 1991) which has been popular in studies of spatial population genetics. In particular,
 376 an extension of the Fleming-Viot process, known as the Λ -Fleming-Viot process, was introduced by Bertoin
 377 and Le Gall (2003) and coined by Etheridge (2008) where it was used to resolve some technical challenges in
 378 modelling isolation by distance (Felsenstein 1975; see also Barton, Etheridge, and Véber 2013; and Barton
 379 and Etheridge 2019). Although this provides an impressive list of accomplishments, the Dawson-Watanabe
 380 superprocess falls short of our needs. In particular this process assumes individuals do not interact and thus
 381 precludes its ability to model ecological interactions. However, this concern has been addressed, leading to
 382 constuctions of superprocesses that account for interactions among individuals. In the next subsection we
 383 summarize the main results in this area and introduce the SPDE that provides the basis for our approach
 384 to theoretical evolutionary ecology.

385 Interacting superprocesses

386 The existence of diffusion limits for a class of measure-valued branching processes involving interactions
 387 among individuals has been treated by Méléard and Roelly (1992, 1993). The interactions can manifest as
 388 dependencies of the spatial movement or reproductive law of individuals on their position x and the state
 389 of the whole population described by X_t . An important result of Méléard and Roelly (1992, 1993) is a
 390 theorem that provides sufficient conditions to construct a sequence of rescaled measure-valued branching
 391 processes that converge to a generalization of the Dawson-Watanabe superprocess that includes interactions.
 392 The rescaling is analogous to that described above for non-interacting Dawson-Watanabe superprocesses,
 393 but now the reproductive law described by $\mathcal{W}(X_t, x)$ and $V(X_t, x)$, branching rate $\lambda(X_t, x)$ and diffusion
 394 parameter $\sqrt{\mu(X_t, x)}$ are allowed to depend on the whole population X_t and individual location x . In Figure
 395 2 we demonstrate this rescaling in discrete time for a population experiencing stabilizing selection and logistic
 396 growth. Since time is discretized, the process we simulate is formally a branching random walk. For further
 397 details on our simulation see SM §5.4.

398 Interactions that manifest in the spatial movement can be used to model mutation bias and those manifesting
 399 in the reproductive law can model density-dependent growth rates and frequency-dependent selection.
 400 Perkins (1992, 1995) developed a theory of stochastic integration with respect to the so-called *Brownian*
 401 *trees* to characterize interacting superprocesses and establish properties of existence and uniqueness. Li
 402 (1998) built directly off of the construction of Méléard and Roelly (1992, 1993) to study properties of density
 403 processes associated with interacting superprocesses, arriving at a SPDE that forms the foundation of our
 404 approach.

405 Recall, we use $\nu(x, t)$ to denote the density of a superprocess, given it exists. Assuming the interactions
 406 manifest only in the reproductive law and that spatial movement follows Brownian motion with diffusion
 407 parameter $\sqrt{\mu}$ independent of both X_t and x , Li (1998) proved a result that implies the interacting super-
 408 process on one dimensional trait space has a density $\nu(x, t)$ which is non-negative, integrable, continuous in
 409 time and space and satisfies the SPDE

$$\frac{\partial}{\partial t}\nu(x, t) = m(\nu, x)\nu(x, t) + \frac{\mu}{2} \frac{\partial^2}{\partial x^2}\nu(x, t) + \sqrt{V\nu(x, t)}\dot{W}(x, t). \quad (25)$$

410 Note that, comparing equation (25) to equation (3.5) of Li (1998), our m and V correspond to Li's b and c
 411 respectively. Denoting $C_1^+(\mathbb{R})$ the set of non-negative, continuous, integrable functions on \mathbb{R} , it is important
 412 to note that, under the assumptions made in Méléard and Roelly (1992, 1993) and Li (1998), equation
 413 (25) is only formal when $m(\nu, x)$ is bounded across all combinations of $\nu \in C_1^+(\mathbb{R})$ and $x \in \mathbb{R}$. However,
 414 recalling our condition $m(\nu, x) \leq r \in \mathbb{R}$, the growth rates we consider are only bounded above. Replacing
 415 $m(\nu, x)$ with its upper-bound demonstrates that the associated superprocess should still admit an integrable
 416 density. That is, we should still be able to compute the total mass process $N(t)$. In fact, in the proof of
 417 the construction of the interacting superprocess as the limit of rescaled branching diffusions, Méléard and
 418 Roelly (1992, 1993) assumed $m(\nu, x)$ to be bounded to guarantee the total mass process will have finite
 419 mean and variance. This allowed them to employ a tightness criterion for sequences of measures and show
 420 the rescaled processes converge to a superprocess with finite total mass. Li's (1998) result builds directly
 421 off of Méléard and Roelly's construction, inheriting the assumption of boundedness for $m(\nu, x)$. However, in
 422 Li (1998), the sufficiency of $m(\nu, x)$ being bounded above is even more clear since Li works explicitly with

Rescalings of a branching random walk

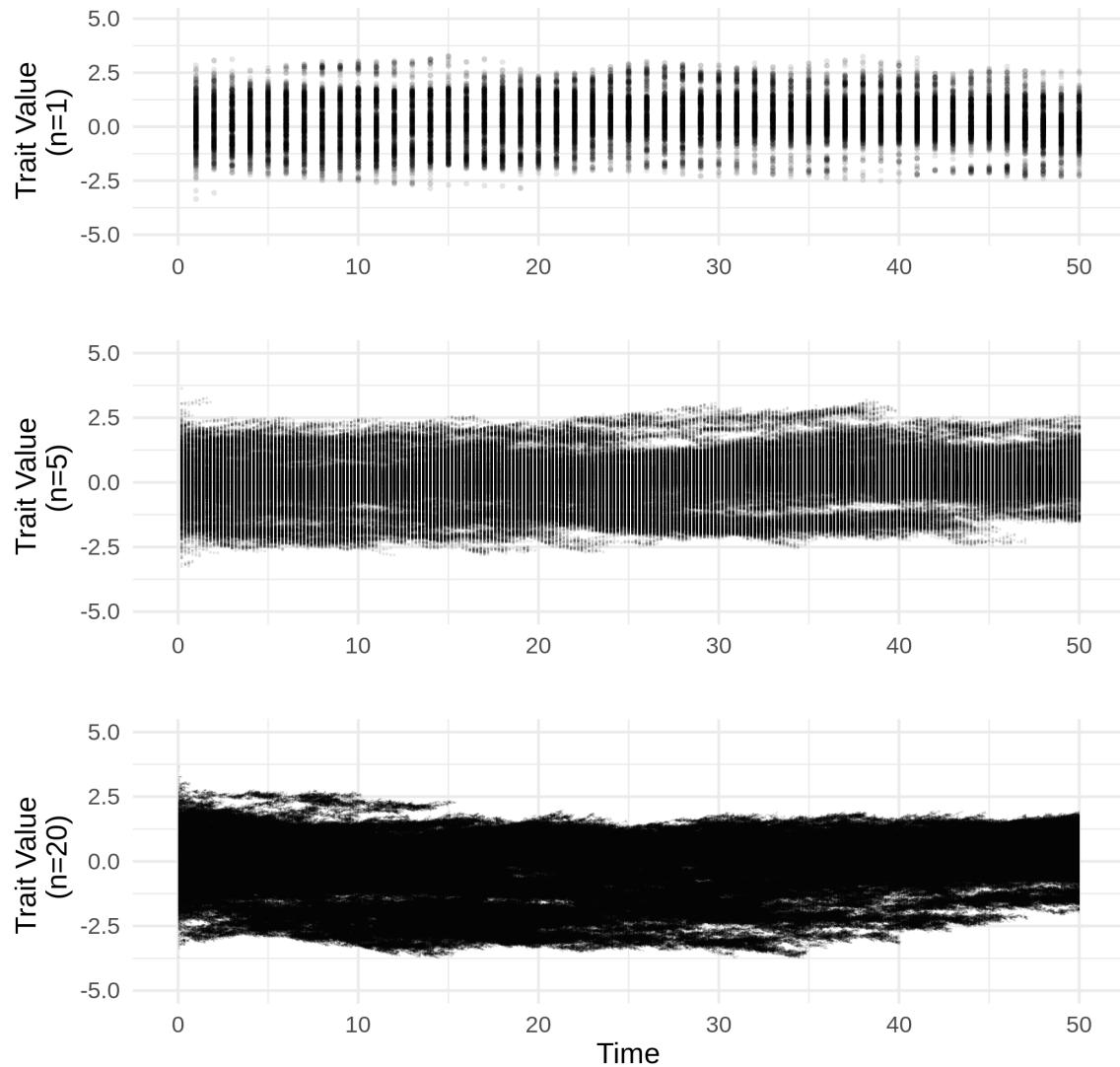


Figure 2: Rescaled sample paths of a branching random walk under stabilizing selection and logistic growth. The top plot displays a sample path without scaling ($n = 1$), the middle plot shows a sample path rescaled by $n = 5$ and the bottom plot shows a sample path rescaled by $n = 20$.

423 a common upperbound for both $m(\nu, x)$ and V . Hence, one can repeat the necessary proofs replacing the
424 assumption that $m(\nu, x)$ is bounded with the assumption that $m(\nu, x)$ is merely bounded above to derive
425 the same results.

426 What remains is to show the spatial mean and variance are finite, assuming finite initial conditions. That is,
427 setting $M_n(t) = \int_{\mathbb{R}} |x|^n \nu(x, t) dx$, we want $M_1(0), M_2(0), M_4(0) < \infty$ to imply $M_1(t), M_2(t), M_4(t) < \infty$ for
428 each $t > 0$ (it turns out we need $M_4(t) < \infty$ for the SDE describing the evolution of phenotypic variance).
429 By condition (??), we see that $m(\nu, x)$ drops quadratically as x leaves θ . In SM §5.1 we show that, when
430 $V = 0$, this form of fitness generates a restoring force that ensures a finite mean and variance. Since mutation
431 causes finite perturbations we do not expect it to overcome this restoring force in the stochastic case. If
432 anything, we suspect a large mutation variance to merely increase the genetic load of the population and
433 drive it to extinction. Indeed, this suggestion is supported by equilibrium of the corresponding deterministic
434 system found in SM §5.4. However, to establish further support, in SM §5.4 we investigate these claims by
435 comparing simulations of the BBM with selection to analytical solutions of the PDE obtained from equation
436 (25) by setting $V = 0$. The results suggest our assertions are sound, but they have yet to be rigorously
437 justified. For now, we leave this as an open problem and assume spatial means and variances are finite for
438 every $t \in [0, \infty)$. In the next subsection we derive an SDE for the abundance of a population and report
439 the SDE for the trait mean \bar{x} and trait variance σ^2 . For the sake of space, derivations of $d\bar{x}$ and $d\sigma^2$ can be
440 found in SM §5.5.

441 2.2.3 Deriving SDE from a SPDE

442 The dynamics of abundance (referred to in the superprocess literature as the total mass process) are obtained
443 by defining the process $N(t) = \int_{\mathbb{R}} \nu(x, t) dx$ and evaluating expression

$$\begin{aligned} & \int_{\mathbb{R}} f(x) \nu(x, t) dx - \int_{\mathbb{R}} f(x) \nu(x, 0) dx \\ &= \int_0^t \int_{\mathbb{R}} \nu(x, s) \left(m(\nu, x) f(x) + \frac{\mu}{2} \frac{\partial^2}{\partial x^2} f(x) \right) dx ds + \int_0^t \int_{\mathbb{R}} f(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) dx ds \end{aligned} \quad (26)$$

444 with $f(x) \equiv 1$ to obtain

$$N(t) - N(0) = \bar{m}(t) N(t) + \int_0^t \int_{\mathbb{R}} \sqrt{V \nu(x, s)} \dot{W}(x, s) dx ds. \quad (27)$$

445 Equation (21) implies

$$\mathbb{V} \left(\int_0^t \int_{\mathbb{R}} \sqrt{V \nu(x, s)} \dot{W}(x, s) dx ds \right) = \int_0^t \int_{\mathbb{R}} V \nu(x, s) dx ds = \int_0^t V N(s) ds. \quad (28)$$

446 Following the notation introduced in §2.2.1, we have

$$\int_0^t d\hat{\mathbf{W}}_s (\sqrt{V \nu(x, s)}) = \int_0^t \int_{\mathbb{R}} \frac{\sqrt{V \nu(x, s)}}{\sqrt{V N(s)}} \dot{W}(x, s) dx ds \quad (29)$$

447 is, as a function of t , a standard Brownian motion. To clean up notation set $dW_1(t) = d\hat{\mathbf{W}}_t (\sqrt{V \nu(x, t)})$.
448 Then we can write an SDE for N which we provide in the following section.

449 Unfortunately, the same trick does not work for calculating SDE for the mean and variance of traits. Indeed,
450 setting $f(x) = x, x^2$ and evaluating expression (26) will lead to SDE that miscalculate the effect of mutation
451 and hence are clearly incorrect results. Obtaining the correct results turns out to require some very involved
452 calculations which we perform in SM §5.5.

453 **2.3 Equations of evolutionary and demographic dynamics**

454 In SM §5.5 we use our assumptions that trait mean and variance are finite and that variance in reproductive
 455 output $V \geq 0$ does not depend on x or $\nu(x, t)$ to show SDE for $N(t)$, $\bar{x}(t)$ and $\sigma^2(t)$ can be expressed as

$$dN(t) = \bar{m}(t)N(t)dt + \sqrt{VN(t)}dW_1(t), \quad (30a)$$

$$d\bar{x}(t) = \text{Cov}_t(x, m(\nu, x))dt + \frac{\mu}{2} \int_{-\infty}^{+\infty} x \frac{\partial^2}{\partial x^2} p(x, t)dx + \sqrt{V \frac{\sigma^2(t)}{N(t)}} dW_2(t), \quad (30b)$$

$$\begin{aligned} d\sigma^2(t) = & \left(\text{Cov}_t((x - \bar{x}(t))^2, m(\nu, x)) + \frac{\mu}{2} \int_{-\infty}^{+\infty} (x - \bar{x}(t))^2 \frac{\partial^2}{\partial x^2} p(x, t)dx - V \frac{\sigma^2(t)}{N(t)} \right) dt \\ & + \sqrt{V \frac{(x - \bar{x}(t))^4 - \sigma^4(t)}{N(t)}} dW_3(t). \end{aligned} \quad (30c)$$

457 where W_1 , W_2 and W_3 are standard Brownian motions. Dividing by dt one can interpret equations (30) as if
 458 they are ordinary differential equations, but this is not technically rigorous since Brownian motion is nowhere
 459 differentiable with respect to time. In SM §5.5 we show that in general W_1 is independent of both W_2 and
 460 W_3 , but W_2 and W_3 covary.

461 There is quite a bit we can learn from expressions (30). Firstly, setting $V = 0$ recovers the deterministic
 462 dynamics derived in §2.1. Alternatively, one can take $N(t) \rightarrow \infty$ to recover the deterministic dynamics for
 463 $\bar{x}(t)$ and $\sigma^2(t)$. Characteristically, we note the effect of demographic stochasticity on abundance grows with
 464 $\sqrt{N(t)}$. Hence, dividing by N , we find the effects of demographic stochasticity on the per-capita growth
 465 rate diminish with increased abundance. Relating the response to demographic stochasticity derived here
 466 to the effect of random genetic drift derived in classic quantitative genetic theory, if we set $\sigma^2(t) = \sigma^2$ and
 467 $N(t) = N$ constant with respect to time, then integrating the stochastic term in equation (30b) over a single
 468 unit of time returns a normally distributed random variable with mean zero and variance equal to $V\sigma^2/N$. In
 469 particular, assuming perfect inheritance, when reproductive variance is unity ($V = 1$) this random variable
 470 coincides with the effect of random genetic drift on the change in mean trait over a single generation derived
 471 using sampling arguments (Lande 1976). There is also an interesting connection with classical population
 472 genetics. A fundamental result from early population genetic theory is the expected reduction in diversity
 473 due to the chance loss of alleles in finite populations (Fisher 1923; Wright 1931). This expected reduction
 474 in diversity due to random genetic drift is captured by the third term in the deterministic component of
 475 expression (30c), particularly $-V\sigma^2(t)/N(t)$. The component of SDE (30c) describing random fluctuations
 476 in $\sigma^2(t)$ is more complicated and is proportional to the root of the difference between the centralized fourth
 477 moment of $p(x, t)$ and $\sigma^4(t)$. These expressions can be used to investigate the dynamics of the mean and
 478 variance for general $\nu(x, t)$. However, in the next subsection we simplify these expressions by approximating
 479 $\nu(x, t)$ with a Gaussian curve. In SM §5.5 we show that under the Gaussian case W_1 , W_2 and W_3 are
 480 independent.

481 **2.3.1 Particular results assuming a Gaussian phenotypic distribution**

482 By assuming $\nu(x, t)$ can be approximated by a Gaussian curve for each $t \geq 0$, expressions (30a), (30b) and
 483 (30b) transform into efficient tools for deriving the dynamics of populations given a fitness function $m(\nu, x)$.
 484 Gaussian phenotypic distributions are often obtained through Gaussian, exponential or weak selection ap-
 485 proximations together with a simplified model of inheritance and random mating (Lande 1980; Turelli 1984,
 486 1986, 2017; Bürger 2000). Alternatively, it has been shown that a Gaussian distribution can provide a rea-
 487 sonable approximation even when selection is strong and non-Gaussian (Turelli and Barton 1994). However,
 488 our approach adds an additional layer of difficulty. Even with Gaussian selection, the resulting solution to
 489 SPDE (25) will only be a Gaussian curve in expectation, assuming a Gaussian initial condition. Yet this
 490 difficulty is not as challenging as it may first appear. Indeed, since SPDE (25) can be derived as a diffusion

491 limit we know that, under the appropriate assumptions on selection, genetic architecture and reproduction,
 492 the stochastic departure from a Gaussian curve is negligible when the ratio V/N is small (i.e., when the
 493 variance in reproductive output is much smaller than the population size). In SM §5.4 we demonstrate this
 494 result using numerical methods. Mathematically, this requirement restricts model parameters to regions
 495 that maintain large population sizes. Biologically, this implies populations are not at risk of extinction.
 496 Hence, models developed in this framework are not suitable for studying colonization-extinction dynamics or
 497 evolutionary rescue. Allowing for these restrictions, we may safely assume that ν is approximately Gaussian
 498 and justify writing

$$\nu(x, t) = \frac{N(t)}{\sqrt{2\pi\sigma^2(t)}} \exp\left(-\frac{(x - \bar{x}(t))^2}{2\sigma^2(t)}\right). \quad (31)$$

499 Under this assumption, Lande (1976) found (suppressing the dependency on t)

$$\text{Cov}(x, m) = \sigma^2 \left(\frac{\partial \bar{m}}{\partial \bar{x}} - \frac{\overline{\partial m}}{\partial \bar{x}} \right). \quad (32)$$

500 Under the same assumption, in SM §5.3 we show

$$\text{Cov}((x - \bar{x})^2, m) = 2\sigma^4 \left(\frac{\partial \bar{m}}{\partial \sigma^2} - \frac{\overline{\partial m}}{\partial \sigma^2} \right) \quad (33)$$

501 and $\overline{(x - \bar{x})^4} = 3\sigma^4$. In particular, this implies

$$d\bar{x} = \sigma^2 \left(\frac{\partial \bar{m}}{\partial \bar{x}} - \frac{\overline{\partial m}}{\partial \bar{x}} \right) dt + \sqrt{V \frac{\sigma^2}{N}} dW_2, \quad (34a)$$

$$d\sigma^2 = 2\sigma^4 \left(\frac{\partial \bar{m}}{\partial \sigma^2} - \frac{\overline{\partial m}}{\partial \sigma^2} \right) dt + \left(\mu - V \frac{\sigma^2}{N} \right) dt + \sigma^2 \sqrt{\frac{2V}{N}} dW_3. \quad (34b)$$

503 These equations allow us to derive the response in trait mean and variance by taking derivatives of fitness,
 504 a much more straightforward operation than calculating a covariance for general phenotypic distributions.
 505 Note that in the above expressions, the partial derivatives of \bar{m} represent frequency independent selection
 506 and the averaged partial derivatives of m represent frequency dependent selection. This relationship has
 507 already been pointed out by Lande (1976) for the evolution of the mean trait, but here we see an analogous
 508 relationship holds also for the evolution of trait variance.

509 In the next subsection we generalize this result to the case when traits are imperfectly inherited. In this
 510 case, the phenotypic variance σ^2 is replaced by a genetic variance G . This genetic variance represents the
 511 component of the variance in expressed traits σ^2 explained by additive effects of different alleles among
 512 genetic loci encoding for the focal phenotype (Roughgarden 1979; Bulmer 1980; Lynch and Walsh 1998). It
 513 is therefore fitting that G is referred to as the additive genetic variance.

514 2.3.2 The evolution of additive genetic variance

515 To model imperfect heritability we consider the relationship between expressed phenotypes $x \in \mathbb{R}$ and
 516 associated genetic values $g \in \mathbb{R}$ known as *breeding values*. The breeding value of an individual is the
 517 sum of additive effects of the alleles carried by the individual on its expressed trait. Since our derivations
 518 of evolutionary equations are based on branching processes that assume asexually reproducing populations
 519 (§2.2.2), the additive genetic variance G is just the variance of breeding values in a population. For a detailed
 520 treatment of breeding values and additive genetic variances, see Bulmer (1980) and Lynch and Walsh (1998).
 521 Following standard assumptions of classical quantitative genetics we assume that the expressed trait for any
 522 given individual is normally distributed around their breeding value with variance η . Hence, $\sigma^2 = G + \eta$.
 523 This coincides with assuming breeding values can be predicted from expressed traits using ordinary least
 524 squares. In the case that all of the effects of alleles on an expressed trait are additive, η is known as the

525 variance of environmental deviation (Lynch and Walsh 1998). For a given breeding value, we denote the
 526 probability density of a randomly drawn expressed trait by $\psi(x, g)$. Hence,

$$\psi(x, g) = \frac{1}{\sqrt{2\pi}\eta} \exp\left(-\frac{(x-g)^2}{2\eta}\right). \quad (35)$$

527 To include this relationship in our framework, we write $\rho(g, t)$ as the abundance density of breeding values
 528 at time t so that $\int_{-\infty}^{+\infty} \rho(g, t) dg = \int_{-\infty}^{+\infty} \nu(x, t) dx = N(t)$. We switch our focus from directly modelling the
 529 evolution of $\nu(x, t)$ to modelling the evolution of $\rho(g, t)$. Once $\rho(g, t)$ is determined, we can compute $\nu(x, t)$
 530 via

$$\nu(x, t) = \int_{-\infty}^{+\infty} \rho(g, t) \psi(x, g) dg. \quad (36)$$

531 However, since selection acts on expressed phenotypes, we use our assumed relationship between breeding
 532 values and expressed traits to calculate the fitness of breeding values. To motivate our approach, consider
 533 the problem of inferring the breeding value of an individual given its expressed trait x . Denote \mathbf{g} a random
 534 variable representing the unknown breeding value. Under our model of inheritance we know x is a random
 535 sample from a normal distribution with mean \mathbf{g} and variance η . Maximizing likelihood suggests x is our best
 536 guess for \mathbf{g} , but the actual value of \mathbf{g} is normally distributed around x with the variance η . Hence, for fixed
 537 x , we obtain $\psi(x, g)$ as the probability density of \mathbf{g} . Thus, the mean fitness of a breeding value g across all
 538 individuals carrying x can be written as

$$m^*(\rho, g) = \int_{-\infty}^{+\infty} m(\nu, x) \psi(x, g) dx. \quad (37)$$

539 This is similar to the approach taken by Kimura and Crow (1978) to calculate the overall effects of selection
 540 for expressed characters onto the changes in the distribution of alleles encoding those characters. However,
 541 instead of focusing on the frequencies of alleles at particular loci, we focus on the densities of breeding
 542 values. With the relationship between $m(\nu, x)$ and $m^*(\rho, g)$ established, we define the evolution of $\rho(g, t)$ by
 543 the SPDE

$$\dot{\rho}(g, t) = m^*(\rho, g) \rho(g, t) + \frac{\mu}{2} \frac{\partial^2}{\partial^2 g} \rho(g, t) + \sqrt{V\rho(g, t)} \dot{W}(g, t). \quad (38)$$

544 We assume $\rho(g, t)$ is Gaussian which implies its mode coincides with \bar{x} . Furthermore, since $\sigma^2 = G + \eta$, we
 545 can use equation (37) and the chain rule from calculus to find

$$\frac{\partial \bar{m}}{\partial G} = \frac{\partial \bar{m}}{\partial \sigma^2} \frac{\partial \sigma^2}{\partial G} = \frac{\partial \bar{m}}{\partial \sigma^2}, \quad (39a)$$

$$\frac{\overline{\partial m}}{\partial G} = \frac{\overline{\partial m}}{\partial \sigma^2} \frac{\partial \sigma^2}{\partial G} = \frac{\overline{\partial m}}{\partial \sigma^2}. \quad (39b)$$

547 Thus, equations (34) become

$$d\bar{x} = G \left(\frac{\partial \bar{m}}{\partial \bar{x}} - \frac{\overline{\partial m}}{\partial \bar{x}} \right) dt + \sqrt{V \frac{G}{N}} dW_2, \quad (40a)$$

$$dG = 2G^2 \left(\frac{\partial \bar{m}}{\partial G} - \frac{\overline{\partial m}}{\partial G} \right) dt + \left(\mu - V \frac{G}{N} \right) dt + G \sqrt{\frac{2V}{N}} dW_3. \quad (40b)$$

549 From expressions (40) we see that, under our model of inheritance, focusing on additive genetic variance G
 550 instead the variance in expressed traits σ^2 makes no structural changes to the basic equations describing the
 551 dynamics of populations.

552 3 A model of diffuse coevolution

553 3.1 Formulation

554 In this section we demonstrate the use of our framework by developing a model of diffuse coevolution across a
 555 guild of S species whose interactions are mediated by resource competition along a single niche axis. Because
 556 our approach treats abundance dynamics and evolutionary dynamics simultaneously, this model allows us
 557 to investigate the relationship between selection gradients and competition coefficients.

558 The dynamics of phenotypic distributions and abundances have been derived above and so the only task
 559 remaining is the formulation of a fitness function. Our approach mirrors closely the theory developed by
 560 MacArthur and Levins (1967), Levins (1968) and MacArthur (1969, 1970, 1972). The most significant
 561 difference, aside from allowing evolution to occur, is the treatment of resource quality, which we replace with
 562 a model of abiotic stabilizing selection. A derivation is provided in SM §5.7.

563 For species i we inherit the above notation for trait value, distribution, average, variance, abundance, etc
 564 except with an i in the subscript. Real world examples of niche axes include the body size of prey for lizard
 565 predators and the date of activity in a season for pollinators competing for floral resources. For mathematical
 566 convenience, we model the axis of resources by the real line \mathbb{R} . The value of a resource along this axis is
 567 denoted by the symbol ζ . For an individual in species i , we assume the resource utilization curve u_i can be
 568 written as

$$u_i(\zeta, x_i) = \frac{U_i}{\sqrt{2\pi w_i}} \exp\left(-\frac{(x_i - \zeta)^2}{2w_i}\right). \quad (41)$$

569 We further assume the niche center x_i is normally distributed among individuals in species i , but the niche
 570 breadth w_i and total niche utilization U_i are constant across individuals in species i and therefore cannot
 571 evolve. Suppose $\theta_i \in \mathbb{R}$ is the optimal location along the niche axis for species i such that, in the absence
 572 of competition, individuals leave on average Q_i offspring when concentrated at θ_i . We capture the rate by
 573 which the fitness falls as niche location ζ leaves the optimum θ_i by the parameter $A_i \geq 0$. Hence, abiotic
 574 stabilizing selection along the resource axis can be modelled by the curve

$$e_i(\zeta) = Q_i \exp\left(-\frac{A_i}{2}(\theta_i - \zeta)^2\right). \quad (42)$$

575 The effect of abiotic stabilizing selection on the fitness for an individual of species i with niche location x_i is
 576 then given by

$$\int_{-\infty}^{+\infty} e_i(\zeta) u_i(\zeta, x_i) d\zeta = \frac{Q_i U_i}{\sqrt{A_i w_i + 1}} \exp\left(-\frac{A_i}{2(A_i w_i + 1)}(\theta_i - x_i)^2\right). \quad (43)$$

577 To determine the potential for competition between individuals with niche locations x_i and x_j , belonging to
 578 species i and j respectively, we compute the niche overlap

$$\mathcal{O}_{ij}(x_i, x_j) = \int_{-\infty}^{+\infty} u_i(\zeta, x_i) u_j(\zeta, x_j) d\zeta = \frac{U_i U_j}{\sqrt{2\pi(w_i + w_j)}} \exp\left(-\frac{(x_i - x_j)^2}{2(w_i + w_j)}\right). \quad (44)$$

579 A notable criticism of using niche overlap to measure the intensity of competition points to cases where
 580 populations competing on multiple niche axes exhibit overlap on at least one of the axes, but no overall
 581 niche overlap (Holt 1987). Thus niche overlap on lower-dimensional projections of some multivariate niche
 582 space does not imply the populations compete. To illustrate with a simple example, consider two populations
 583 competing for space on the plane \mathbb{R}^2 . If the spatial distributions of the two populations overlap, then
 584 they will overlap on both spatial axes. However, if the populations do not overlap on at least one of the

585 spatial axes, they will have no overall spatial overlap. Furthermore, even if the species overlap on both
 586 spatial axes, they need not have any overall spatial overlap. This final result corresponds to the fact that
 587 components of niche space do not necessarily interact multiplicatively to determine the consequences for
 588 the intensity of competition. In another component of Holt's (1987) critique, an argument is made for the
 589 potential of competition occurring without any overlap in niche space. However, this argument is based on
 590 the practical difficulty of identifying every resource axis populations are competing on and how these axes
 591 interact to determine fitness consequences. Our model avoids these caveats by assuming competition only
 592 occurs along a single dimensional resource gradient. To map the degree of niche overlap to fitness, we assume
 593 competition between individuals with niche locations x_i and x_j additively decreases the Malthusian fitness
 594 for the individual in species i by $c_i \mathcal{O}_{ij}(x_i, x_j)$ for some $c_i \geq 0$. In SM §5.7 we combine this niche model with
 595 equations (30a), (40a) and (40b) to find

$$dN_i = \left\{ R_i - \frac{a_i}{2} ((\bar{x}_i - \theta_i)^2 + G_i + \eta_i) - c_i \sum_{j=1}^S N_j U_i U_j \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2} \right\} N_i dt + \sqrt{V_i N_i} dW_1, \quad (45a)$$

$$d\bar{x}_i = \left\{ a_i G_i (\theta_i - \bar{x}_i) - c_i G_i \left(\sum_{j=1}^S N_j U_i U_j b_{ij} (\bar{x}_j - \bar{x}_i) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2} \right) \right\} dt + \sqrt{V_i \frac{G_i}{N_i}} dW_2, \quad (45b)$$

$$\begin{aligned} dG_i = & \left\{ c_i G_i^2 \left(N_i U_i^2 b_{ii} \sqrt{\frac{b_{ii}}{2\pi}} + \sum_{j=1}^S N_j U_i U_j b_{ij} (1 - b_{ij}(\bar{x}_i - \bar{x}_j)^2) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2} \right) \right. \\ & \left. + \mu_i - a_i G_i^2 - V_i \frac{G_i}{N_i} \right\} dt + G_i \sqrt{\frac{2V_i}{N_i}} dW_3, \end{aligned} \quad (45c)$$

596 where

$$R_i = \ln \left(\frac{Q_i U_i}{\sqrt{1 + A_i w_i}} \right), \quad (46a)$$

$$a_i = \frac{A_i}{1 + A_i w_i}, \quad (46b)$$

$$b_{ij}(t) = b_{ji}(t) = (w_i + w_j + \eta_i + \eta_j + G_i(t) + G_j(t))^{-1}, \quad (46c)$$

$$c_i \geq 0. \quad (46d)$$

597 Despite the convoluted appearance of system (45), there are some nice features that reflect biological reasoning.
 598 For example, the dynamics of abundance are just a generalization of Lotka-Volterra dynamics. In
 599 particular, the effect of competition with species j on the fitness of species i grows linearly with N_j . However,
 600 as biotic selection pushes \bar{x}_i away from \bar{x}_j , the effect of competition with species j on the fitness of species i
 601 rapidly diminishes, reflecting a reduction in niche overlap. The divergence of \bar{x}_i and \bar{x}_j due to competition
 602 is referred to in the community ecology literature as character displacement. We also see that the fitness
 603 of species i drops quadratically with the difference between \bar{x}_i and the abiotic optimum θ_i . Hence, abiotic
 604 selection acts to pull \bar{x}_i towards θ_i . The response in mean trait \bar{x}_i to natural selection is proportional to the
 605 amount of heritable variation in the population, represented by the additive genetic variance G_i . However,
 606 we have that G_i is itself a dynamic quantity. Under our model, abiotic stabilizing selection erodes away
 607 heritable variation at a rate that is independent of both N_i and \bar{x}_i . The effect of competition on G_i is a
 608 bit more complicated. When $b_{ij}(\bar{x}_i - \bar{x}_j)^2 < 1$, competition with species j acts as diversifying selection
 609 which tends to increase the amount of heritable variation. However, when $b_{ij}(\bar{x}_i - \bar{x}_j)^2 > 1$, competition
 610 with species j acts as directional selection and reduces G_i . In the following subsections we demonstrate
 611 the behavior of system (45) by plotting numerical solutions and investigate implications for the relationship
 612 between the strength of ecological interactions and selection.

613 3.2 Community dynamics

614 For the sake of illustration we numerically integrated system (45) for a richness of $S = 100$ species. We
615 assumed homogeneous model parameters across species in the community as summarized by Table 2. We
616 repeated numerical integration under the two scenarios of weak and strong competition. For the first scenario
617 of weak competition we set $c = 1.0 \times 10^{-7}$ and for the second scenario of strong competition we set $c =$
618 5.0×10^{-6} . With these two sets of model parameters, we simulated our model for 1000.0 units of time.
619 For both scenarios, we initialized the trait means to $\bar{x}_i = 0.0$, additive genetic variances to $G_i = 10.0$ and
620 abundances to $N_i = 1000.0$ for each $i = 1, \dots, S$.

621 Temporal dynamics for each scenario are provided in Figure 3. This figure suggests weaker competition leads
622 to smoother dynamics and a higher degree of organization within the community. Considering expression
623 (45a) we note that, all else equal, relaxed competition allows for larger growth rates which promote greater
624 abundances. From (45a) we also note that the per-capita effects on demographic stochasticity diminish with
625 abundance. To see this, divide both sides by N_i . Inspecting expressions (45b) and (45c), we see that larger
626 abundances also erode the effects of demographic stochasticity on the evolution of mean trait and additive
627 genetic variance. These effects were already noted in §2.3, and thus are not a consequence of our model of
628 coevolution per-se, but we revisit them here since Figure 3 demonstrates the importance of demographic
629 stochasticity in structuring ecological communities even when populations are very large. Hence, contrary
630 to the common assumption that stochastic effects can be ignored for large populations, we find that minute
631 asymmetries by demographic stochasticity remain significant drivers of community structure. In particular,
632 we initialized the species with identical state variables and model parameters, but found an enormous
633 amount of asymmetry and even some potential phase changes. In the following two paragraphs we describe
634 the natural history of the community as illustrated in Figure 3.

635 After about 125.0 units of time, the community appears to have shaken off the initial conditions and entered
636 into a qualitatively distinct phase of dynamics. Aside from a few outliers, most of the species remain clustered
637 together in their state variables. This lasts for approximately 375.0 units of time until, at around time 500.0,
638 a drastic change occurs. At this moment the tightly packed cluster of species begins to fan out in all three
639 state variables. Simultaneously, we observe large a shift in mean traits for higher values and in additive
640 genetic variances for lower values. Upon inspecting our calculations, we diagnose the reason for this shift.
641 The outlier species that were initially pushed away from the common abiotic optimum (0.0 in this case)
642 evolved a significant reduction in the quantity of heritable variation ($\approx 60\%$) due to directional selection
643 induced by competition. This reduction in heritable variation slowed adaptation, causing these species to
644 linger on the outskirts of niche space, some longer than others. In the meantime the rest of the community,
645 being tightly packed, experienced greater competition which led to diminished abundances for these species
646 and caused some members of the core group to veer away from the abiotic optimum. The reduced abundances
647 of the core group led to reduced competition overall. As a result, the outlier populations were given a slight
648 increase in growth rate, enough to allow them to increase their abundances orders of magnitude higher than
649 the species in the core group and giving them more weight in driving the evolution of other species. Many of
650 these heavy-hitting outlier species had already been maintaining negative mean traits, but around time 500.0
651 the high abundance species with positive mean traits began to experience enough intraspecific competition
652 to override interspecific competition and begin evolving towards the abiotic optimum. The sudden imbalance
653 of these high abundance species effectively induced a single large competitive exclusion event pushing the
654 majority of the community far away from the abiotic optimum. After this shift the cluster began to slowly
655 bloom in all three state variables as species took advantage of novel asymmetries in their competitive abilities
656 mediated by a new distribution of mean trait values across the community. About 125.0 units of time later,
657 the community reached a qualitatively new phase of dynamics. If we kept running the numerical integrator,
658 we would continue to see similar drama unfolding over and over again as minute stochastic changes contribute
659 to asymmetries which slowly build into drastic shifts.

660 The strong competition scenario is not quite as showy. Although the dynamics of trait means and variances
661 tend to be far more stochastic than in the weak competition scenario, the community overall appears to
662 quickly reach some statistical equilibrium and remain there. However, the abundances across all species in
663 the community are very low due to strength of competition being an order of magnitude higher than in the
664 weak competition case. Most of the species maintain abundances greater than 1000.0, but we found one

Table 2: Values of model parameters used for numerical integration.

Parameter	Description	Value
R	innate growth rate, see §3.3	1.0
θ	abiotic optimum	0.0
a	strength of abiotic selection	0.01
c	strength of competition	$\{1.0 \times 10^{-7}, 5.0 \times 10^{-6}\}$
w	niche breadth	0.1
U	total niche use	1.0
η	segregation variance	1.0
μ	mutation rate	1.0×10^{-7}
V	variance of reproductive output	5.0

665 species that dropped to an abundance of about 50.0. If we let the numerical integrator run long enough in
 666 this case, we will likely see many of the species go extinct.

667 Finding ways to interpret simulated dynamics provides a useful arena to exercise biological reasoning. How-
 668 ever, it does not fulfill our desire to quantify the patterns and processes present in competing communities.
 669 In the next subsection we take a step in this direction by using our model to derive formula for selection
 670 gradients and competition coefficients. To investigate their relationship, we calculate their covariances using
 671 simplifying assumptions on species abundances and intraspecific trait variances. We then investigate how
 672 these covariances change with the ratio of variance of interspecific mean traits to variance of intraspecific
 673 individual traits and use a numerical approach to investigate correlations between the strength of pairwise
 674 coevolution and competition coefficients.

675 3.3 The relation between the strength of ecological interactions and coevolution

676

677 Relating our treatment of the niche to modern coexistence theory (Chesson 2000), the absolute competition
 678 coefficient α_{ij} becomes a dynamical quantity that can be written as

$$\alpha_{ij}(t) = \frac{c_i}{r_i(t)} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p_i(x, t) p_j(y, t) \mathcal{O}_{ij}(x, y) dx dy = \frac{c_i U_i U_j}{r_i(t)} \sqrt{\frac{b_{ij}(t)}{2\pi}} \exp\left(-\frac{b_{ij}(t)}{2} (\bar{x}_i(t) - \bar{x}_j(t))^2\right), \quad (47)$$

679 where

$$r_i(t) = R_i - \frac{a_i}{2} ((\bar{x}_i(t) - \theta_i)^2 + G_i(t) + \eta_i). \quad (48)$$

680 Hence, $dN_i(t)$ can be expressed as

$$dN_i(t) = r_i(t) \left(1 - \sum_{j=1}^S \alpha_{ij}(t) N_j(t) \right) N_i(t) dt + \sqrt{V_i N_i(t)} dW_1(t). \quad (49)$$

681 Note that although $r_i(t)$ is referred to in the coexistence literature as the intrinsic growth rate of the
 682 population, R_i is a deeper intrinsic quantity. For now we refer to R_i as the *innate* growth rate. With this
 683 connection formally established, researchers may pursue a postmodern coexistence theory that naturally
 684 includes the evolutionary dynamics of populations and the effects of demographic stochasticity.

685 In SM §5.7 we show that the standardized directional selection gradient (sensu Lande and Arnold 1983)
 686 induced by species j on species i can be computed as

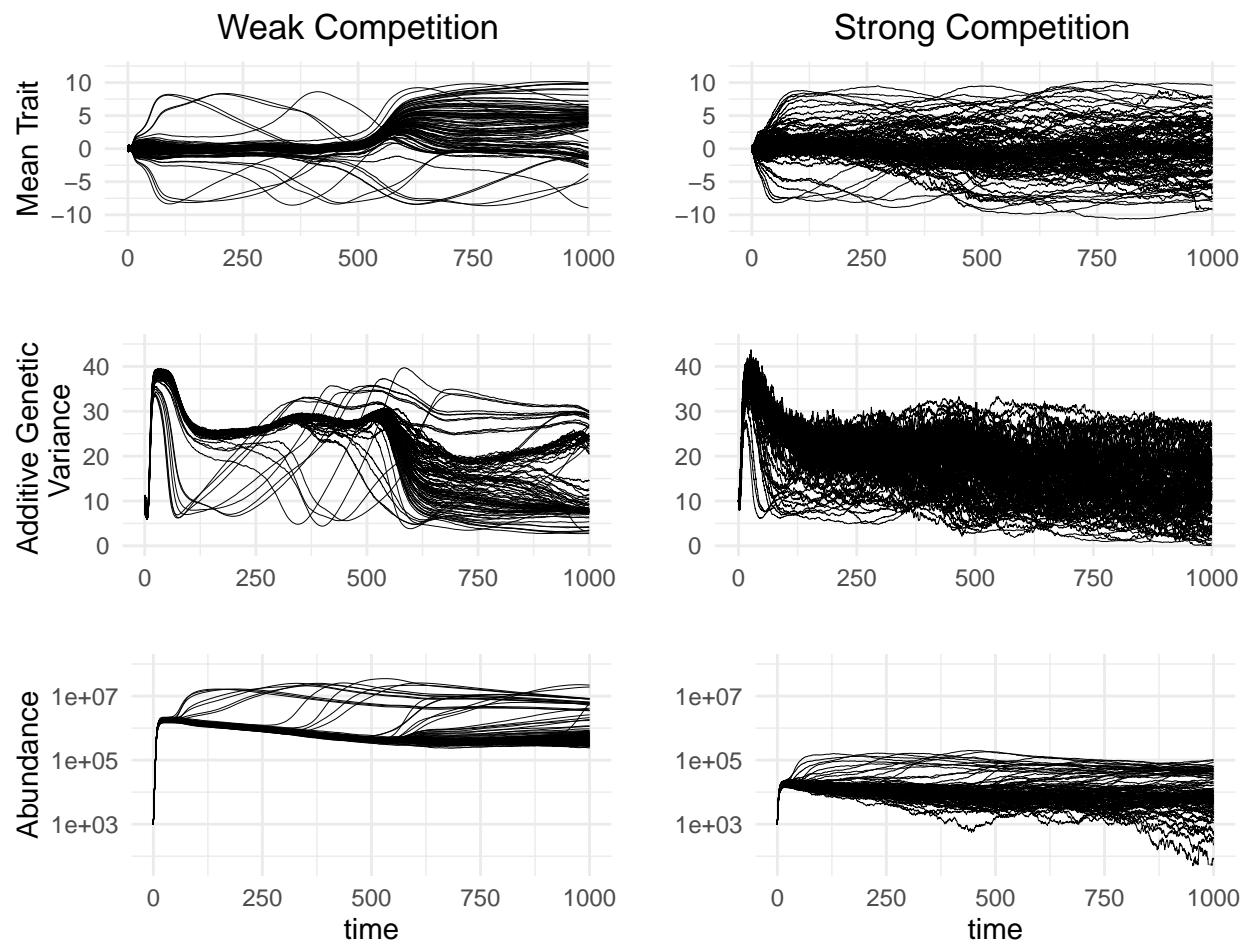


Figure 3: Temporal dynamics of mean trait (top), additive genetic variance (middle) and abundance (bottom) for the scenario of weak competition (left) and strong competition (right). Red lines indicate average trend across species.

$$\beta_{ij}(t) = c_i U_i U_j N_j(t) b_{ij}(t) (\bar{x}_i(t) - \bar{x}_j(t)) \sqrt{\frac{b_{ij}(t)}{2\pi}} \exp\left(-\frac{b_{ij}(t)}{2} (\bar{x}_i(t) - \bar{x}_j(t))^2\right). \quad (50)$$

687 Our notation differs from Lande and Arnold (1983) in that subscripts here denote species instead of components
 688 of multivariate traits and we drop the prime that distinguishes between selection gradients and
 689 standardized selection gradients.

690 Below we investigate the correspondence of interaction intensity and coevolutionary change. However, we
 691 can already identify one major discrepancy; α_{ij} is maximized when $\bar{x}_i = \bar{x}_j$, but $\beta_{ij} = 0$ under the same
 692 condition. We therefore include in our metric of selection the standardized stabilizing selection gradient γ
 693 which measures the effect of stabilizing or disruptive selection on phenotypic variance (Lande and Arnold
 694 1983). In SM §5.7 we show that the standardized stabilizing selection gradient induced by species j on
 695 species i can be computed as

$$\gamma_{ij}(t) = c_i U_i U_j N_j(t) b_{ij}(t) \left(1 - b_{ij}(t) (\bar{x}_i(t) - \bar{x}_j(t))^2\right) \sqrt{\frac{b_{ij}(t)}{2\pi}} \exp\left(-\frac{b_{ij}(t)}{2} (\bar{x}_i(t) - \bar{x}_j(t))^2\right). \quad (51)$$

696 To measure the total evolutionary change in species i induced by species j , we form the metric $\Psi_{ij} =$
 697 $|\beta_{ij}| + |\gamma_{ij}|$. The top row of Figure 4 displays interaction networks under weak and strong competition where
 698 the edge width connecting species i and j is proportional to $\alpha_{ij}\alpha_{ji}$. The bottom row of Figure 4 displays
 699 the distributions of pairwise coevolutionary change, which we measure for species i and j via $\mathfrak{C}_{ij} = \Psi_{ij}\Psi_{ji}$,
 700 under weak and strong competition.

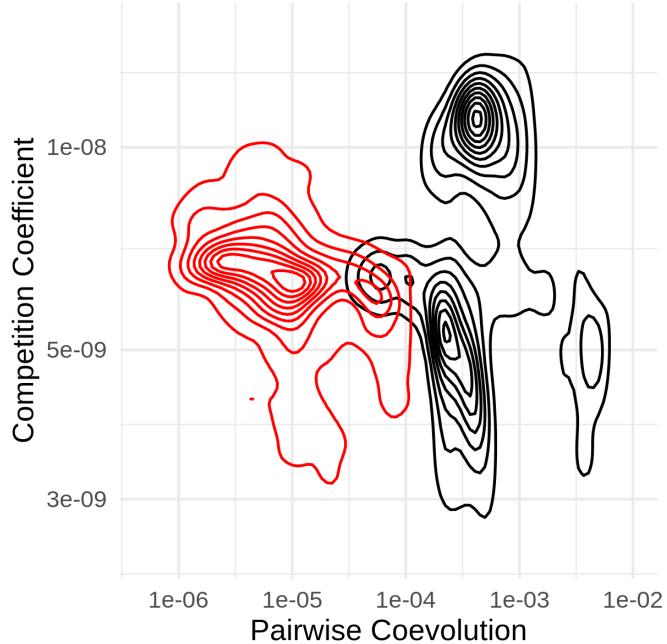


Figure 4: Networks of interspecific interactions parameterized by competition coefficients (top row) and distributions of pairwise coevolution (bottom row) under the scenarios of weak competition (left) and strong competition (right) at time $t = 1.0 \times 10^3$. Node sizes are proportional to population sizes. Edge widths and shade are monotonically increasing functions of pairwise coevolution.

701 We now make use of the expressions derived for competition coefficients and selection gradients to investigate
 702 their relationship. As a first pass, let us assume all model parameters are equivalent across species and that
 703 each species has the same abundance and trait variance. Let us further assume that richness S is large and

704 the distribution of mean trait values is normal with mean \bar{x} , variance $V_{\bar{X}}$ and density $f_{\bar{X}}$. Such assumptions
 705 are typical when deriving analytical results in the field of theoretical coevolutionary community ecology
 706 (Guimarães, Jordano, and Thompson 2011; Guimarães et al. 2017; Nuismer, Jordano, and Bascompte 2012;
 707 Nuismer, Week, and Aizen 2018). If \bar{x} is near θ and $V_{\bar{X}}$ is much smaller than $|2R/a - G - \eta|$, then we may
 708 approximate r_i with

$$\bar{r} = \int_{-\infty}^{+\infty} \left(R - \frac{a}{2}((\bar{x} - \theta)^2 + G + \eta) \right) f_{\bar{X}}(\bar{x}) d\bar{x} = R - \frac{a}{2}((\bar{x} - \theta)^2 + V_{\bar{X}} + G + \eta). \quad (52)$$

709 In SM §5.8 we use these assumptions to calculate the first and second order moments describing the joint
 710 distribution of competition coefficients and selection gradients across the community. We find that the
 711 covariance between linear selection gradients and competition coefficients are zero due to the symmetry
 712 implied by our assumptions. However, setting $\alpha(\bar{x}_i, \bar{x}_j) = \alpha_{ij}$, $\beta(\bar{x}_i, \bar{x}_j) = \beta_{ij}$ and $\gamma(\bar{x}_i, \bar{x}_j) = \gamma_{ij}$, the
 713 covariances between the magnitude of linear selection gradients and competition coefficients and between
 714 stabilizing selection gradients and competition coefficients can be written as

$$\text{Cov}_{f_{\bar{X}}}(\alpha, |\beta|) = \frac{2c^2 b^2 U^4 N}{\pi \bar{r}} \sqrt{\frac{V_{\bar{X}}}{2\pi}} \left(\frac{1}{(1 + 8bV_{\bar{X}})^{3/4}} - \frac{1}{(1 + 4bV_{\bar{X}})^{3/4}(1 + 2bV_{\bar{X}})^{1/2}} \right), \quad (53a)$$

$$\text{Cov}_{f_{\bar{X}}}(\alpha, \gamma) = \frac{c^2 b^2 U^4 N}{2\pi \bar{r}} (1 - 2bV_{\bar{X}}) \left(\frac{1}{\sqrt{1 + 4bV_{\bar{X}}}} - \frac{1}{1 + 2bV_{\bar{X}}} \right). \quad (53b)$$

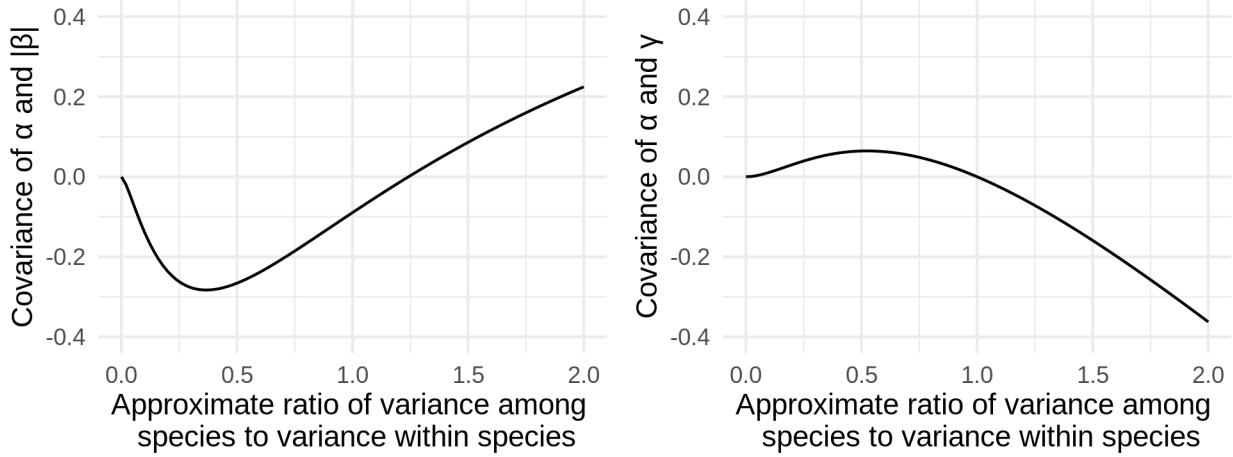


Figure 5: Curves representing the covariance between the magnitude of linear selection gradients and competition coefficients (left) and between stabilizing selection gradients and competition coefficients (right) as a function of $2bV_{\bar{X}}$ which is approximately equal to the ratio of variance in mean traits among species to the intraspecific trait variance. In both plots we set $c = 1.0 \times 10^{-4}$, $b = 0.1$, $\bar{r} = 0.1$ and $N = 1.0 \times 10^{10}$ and let $V_{\bar{X}}$ vary.

715 For fixed c, b, \bar{r} and N , we visualize these relationships in Figure 5. To gain insight into the relationship
 716 between selection gradients and competition coefficients, note that our assumptions in this section imply
 717 $b^{-1} = 2(\sigma^2 + w)$. If we further assume $\sigma^2 + w \approx \sigma^2$, then $2bV_{\bar{X}} \approx V_{\bar{X}}/\sigma^2$. That is, when populations are
 718 generalists and are comprised of specialist individuals, the value $2bV_{\bar{X}}$ is approximately equal to the ratio
 719 of interspecific mean trait variation to intraspecific individual trait variation. Hence, for both covariances
 720 we see that there is no relationship between selection gradients and competition coefficients when this
 721 ratio is zero. From equation (53a) we can use numerical optimization to find that when $V_{\bar{X}}/\sigma^2 \approx 1.25$ the
 722 relationship between the magnitudes of linear selection gradients and competition coefficients disappears, but

723 when (approximately) $V_{\bar{X}}/\sigma^2 < 1.25 (> 1.25)$, this covariance becomes negative (positive). Equation (53b)
724 states that when $V_{\bar{X}}/\sigma^2$ is approximately equal to one (or slightly larger), there is no expected relationship
725 between competition coefficients and quadratic selection gradients. However, when $V_{\bar{X}}/\sigma^2 < 1.0 (> 1.0)$,
726 then we expect a positive (negative) relationship between α and γ . These results are true regardless of the
727 chosen parameter values. In SM §5.8 we use simulations of system (45) to show that these results do not
728 qualitatively differ when allowing for heterogeneous population sizes and additive genetic variances across
729 species.

730 From a biological perspective, if the ratio $V_{\bar{X}}/\sigma^2$ is small, then species are packed tightly in phenotypic space.
731 In our model this occurs when abiotic stabilizing selection is much stronger than competition ($a \gg c$). This
732 causes species to overlap more in niche space (i.e., large α) and creates disruptive selection for greater
733 intraspecific variance (i.e., positive γ), which explains the positive region of covariance between α and γ .
734 However, as species begin to overlap in niche space, directional selection begins to vanish (i.e., small $|\beta|$),
735 leading to a negative covariance between α and $|\beta|$. In the limiting case that two species have perfectly
736 overlapping niches, they will exhibit zero directional selection since a shift in either direction will yield the
737 same fitness advantages.

738 In the opposite scenario where competition is much stronger than abiotic stabilizing selection ($c \gg a$), species
739 will not evolve to be as tightly packed and instead their niche-centers will be spread out with little overlap in
740 their resource utilization curves (i.e., small α). In this case biotic directional selection will be strong (i.e., large
741 $|\beta|$), particularly for species towards the outer regions of niche space due to asymmetric fitness advantages
742 conferred by shifts in niche-centers. This leads to a positive covariance between α and $|\beta|$. However, as
743 noted above, this directional selection will also erode away at standing heritable variation (i.e., negative γ),
744 reducing the rate at which adaptation can occur and creating a negative covariance between α and γ .

745 In summary, we see the relation between competition coefficients and selection is highly non-trivial and
746 depends on the relative magnitudes of different ecological processes shaping the community. However,
747 this does not address the relation between competition coefficients and coevolution per se. In SM §5.8
748 we show that calculating a formula for the covariance between competition coefficients and the metric
749 of coevolution \mathfrak{C} introduced above provides a difficult analytical challenge. Instead of confronting this
750 challenge we build on our numerical approach used to justify analytical approximations of $\text{Cov}_{f_{\bar{X}}}(\alpha, |\beta|)$ and
751 $\text{Cov}_{f_{\bar{X}}}(\alpha, \gamma)$ to approximate the correlation of α and \mathfrak{C} . This numerical approach inherits the assumptions
752 of homogeneous background parameters such as the mutation rate μ and abiotic optima θ , but allows us to
753 relax the assumption that N and G are constant across species and time.

754 In particular, we numerically integrated system (45) for $T_1 = 1000.0$ units of time and then continued to
755 integrate for $T_2 = 1000.0$ units of time. We then calculate the covariance between the quantities α and
756 \mathfrak{C} across $S = 100$ species for each of the last T_2 time steps. We assume the temporal average of these
757 covariances across the last T_2 units of time approximates the expectation at equilibrium. We repeated this
758 approach for randomly drawn a and c until our sample size reached 1000. In Figure 6 we plot the temporally
759 averaged values of $\text{Cov}_{f_{\bar{X}}}(\alpha, \mathfrak{C})$ against the strength of competition c . Using a cubic regression, we see the
760 correlation of coevolutionary selection gradients and competition coefficients is negative at variance ratios
761 below 0.5, zero at variance ratios between 0.5 and 1.0, and drops below zero again at variance ratios above
762 1.0.

763 4 Conclusion

764 We have introduced an approach to derive models of evolutionary ecology using the calculus of white noise,
765 demonstrated our approach by deriving a model of diffuse coevolution and investigated the relationship
766 between selection gradients and competition coefficients, finding that these quantities exhibit interesting
767 relationships which shed light on the feedback between the structure and dynamics of ecological communities.

768 Our approach has the merit of rigorously synthesizing the dynamics of abundance and phenotypic evolution
769 for populations experiencing demographic stochasticity. Yet, there remains biological details and their

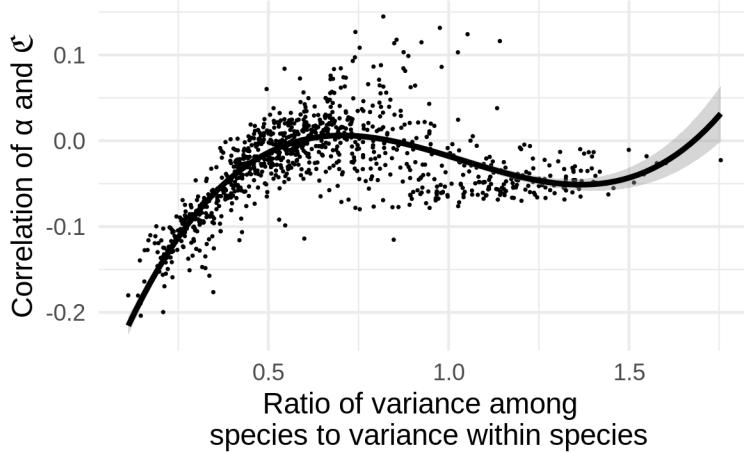


Figure 6: Numerical estimates for the correlation of the strength of coevolution measured by \mathfrak{C} and competition coefficients α plotted against the variance of mean traits among species divided by the mean variance of traits within species. Each dot represents the result from a single simulation. The red line is a cubic regression.

770 associated technical challenges that need to be confronted for gaining a more thorough and rigorous understanding of ecological communities. We touch on just four of them here and provide suggestions for future
771 research to approach these challenges.
772

773 Limitations of diffusion limits

774 As noted by Feller (1951), diffusion limits provide reasonable approximations for large populations, but relatively
775 small populations require discrete models. Hence, as a diffusion limit, SPDE (25) restricts parameter
776 values to regions that maintain large population sizes. Although this suggests an important restriction on
777 any model developed under this framework by implying populations are not at risk of extinction, the SDE
778 describing abundance dynamics technically permits extinction. However, for small abundances, pathology
779 emerges in the evolution of trait means and variances. In particular, stochastic components of the SDE
780 describing the evolution of \bar{x} and σ^2 diverge towards infinite values as $N \rightarrow 0$. Hence, studies of evolutionary
781 rescue and colonization-extinction dynamics that incorporate phenotypic evolution should be pursued
782 via a different approach. A natural alternative can be developed utilizing the underlying BBM that SPDE
783 (25) is a diffusion limit of (see section 2.2.2). This process explicitly tracks individuals throughout their
784 life-history and captures reproduction as branching events. Hence, BBM processes model population size as
785 a non-negative integer instead of a continuously varying number. In particular, the pathological behavior
786 described for the diffusion limit does not occur for BBM as population size goes to zero.

787 The genetic architecture and distributions of traits and the role of sexual reproduction

788 Our treatment of inheritance and our approach to model coevolution rest on the assumptions of normally
789 distributed breeding values and expressed phenotypes along with asexual reproduction. However, real traits
790 are not encoded by an infinite number of loci each contributing an additive infinitesimal effect (as required
791 by the infinitesimal model), mutations are not inherited as normally distributed deviations from parental
792 breeding values (as required by the Gaussian descendants model) and many populations of interest exhibit
793 non-random sexual reproduction. Departures from this model of genetic architecture can produce more
794 complex distributions of breeding values and expressed traits. Such deviations can be reinforced via strong
795 non-Gaussian selection surfaces, including the surface $m(\nu, x)$ we have derived from niche theory, along with
796 non-random mating in sexually reproducing populations. However, Gaussian approximations are convenient
797 since they are defined by their mean and variance. Future work investigating the effects of non-normally dis-
798 tributed traits on the structure and dynamics of ecological communities will therefore need to confront higher
799 moments such as skew and kurtosis, ideally integrating previously established mathematical approaches to

800 derive the dynamics of these higher moments (Débarre, Yeaman, and Guillaume 2015).

801 An alternative approach to breaking the assumption of normally distributed trait values is the development
802 of explicit multilocus models. These models describe the contributions of alleles at particular loci in the
803 genome to the development of quantitative traits. Tracking the fluctuations of allele frequencies then allows
804 theoreticians to investigate the dynamics of phenotypic distributions that deviate from normality. This
805 approach has a long history in theoretical quantitative genetics (Bulmer 1980; Turelli and Barton 1994;
806 Kirkpatrick, Johnson, and Barton 2002) and coevolutionary theory (Nuismer, Doebeli, and Browning 2005;
807 Kopp and Gavrilets 2006; Nuismer, Ridenhour, and Oswald 2007). However, work to investigate the role
808 of genetic architecture in mediating feedbacks between the dynamics of population abundances and the
809 distributions of traits mediating ecological interactions has apparently only just begun (Schreiber, Patel,
810 and terHorst 2018; Patel and Bürger 2019).

811 **The role of ecological stoichiometry**

812 Our treatment of both biotic and abiotic selection neglects important chemical constraints of biological
813 reality. For instance, the resource we assume species are competing over is modelled as a static quantity.
814 However, real resources can be dynamic quantities. Such theoretical quantities may reflect abiotic cycles
815 of material and energy or whole trophic layers comprised of prey, hosts or mutualists. Although resource
816 dynamics have been captured theoretically in consumer-resource models (Tilman 1982), developing a more
817 realistic model of resource competition must incorporate details on the ecophysiology of organisms (Loreau
818 2010). Doing so will help clarify the feedback between the evolution of populations and the ecosystem
819 processes they are a part of.

820 Using plant-pollinator communities as an example, the role of nitrogen mediating interspecific interactions
821 has been reviewed by David, Storkey and Stevens (2019) and the evolutionary ecology of the nutritional
822 content of nectar has been reviewed by Parachnowitsch, Manson and Sletvold (2018). These studies demon-
823 strate the need for further research to understand how soil nutrient availability and organismal ecophysiology
824 structures communities of plants and pollinators. Theoretical pursuits in this directions may profit from in-
825 terfacing the framework we have outlined here with population-ecosystem models such as that developed by
826 Fridley (2017).

827 **Accounting for macroevolutionary history**

828 To understand patterns found in ecological communities, considerations must push beyond microevolutionary
829 and contemporary ecological processes and consider the macroevolutionary dynamics of ancestral lineages
830 leading to extant populations. Using sub-alpine flower communities as an example, one can observe a
831 very strict ordering of phenology across broad geographic ranges. In particular, whether in the Colorado
832 Rocky mountains (such as Gothic, Colorado) or on an outlier of the Idaho batholith (such as Kamiak butte
833 near Palouse, Washington), one will almost surely observe a very conspicuous order of flowers emerging
834 in early spring: at the very beginning of the season blooms *Claytonia lanceolata* followed by *Erythronium*
835 *grandiflorum* which precedes *Delphinium nuttallianum* (B. Week, personal observations). If contemporary
836 phenological coevolution is rampant, why should this pattern be so well preserved across a thousand miles of
837 rugged and diverse terrain? Although it would be exciting to find that these species repeatedly coevolved this
838 pattern in each location, a more parsimonious hypothesis suggests the phenology and physiology of these
839 species slowly evolved independently over macroevolutionary time scales to take advantage of the specific
840 conditions available within each of these windows of the flowering season. However, this could not have
841 carried out in the Rocky mountains since this terrain only became habitable just over ten thousand years
842 ago as the glaciers of the Pleistocene began to retreat (Paul CaraDonna, personal communications). Hence,
843 given these considerations, it appears that an understanding of early season phenology patterns should focus
844 on how these communities are assembled as opposed to contemporary evolutionary dynamics. Indeed, recent
845 work testing models of phylogeography ignores the potential for contemporary evolution and instead suggests
846 alpine flower communities tend to follow neutral assembly where flowers merely compete for who can disperse
847 to new habitat first, as opposed to a selective process where a regional species pool is filtered for those species
848 adapted to the newly available habitat (Marx et al. 2017).

849 Of course microevolutionary and ecological dynamics are not completely irrelevant for understanding patterns
850 in communities that are primarily structured by deep evolutionary processes. In particular, macroevolution-

851 trait evolution is simply the aggregation of microevolutionary change occurring over large spans of time.
852 This suggests a road forward to connect the theory we have introduced to models of macroevolutionary trait
853 evolution.

854 Some approaches to modelling macroevolutionary trait change simply repurpose microevolutionary models
855 by blindly rescaling time from the units of generations to millions of years [Nuismer and Harmon (2014);
856 Luke, can you think of others?]. Such an approach makes the implicit assumption that trait evolution is
857 statistically self-similar (*sensu* Falconer 2014) so that the stochastic evolution of traits on macroevolutionary
858 time scales has the same properties of trait evolution on microevolutionary time scales. Although some
859 stochastic processes, including Brownian motion, do exhibit self-similarity, others do not. For example,
860 consider a modification of the Ornstein-Uhlenbeck process defined by the SDE

$$dX_t = a(\theta_t - X_t)dt + bdW_t \quad (54)$$

861 where $a, b > 0$, W_t is a standard Brownian motion and θ_t is itself a Markov process that takes normally
862 distributed jumps centered on its current location at exponentially distributed time intervals. If we assume
863 the rate λ at which jumps occur is much smaller than a and the variance in jumping is much larger than
864 b^2 , then, even though the sample paths of X_t are actually continuous (if we zoom in close enough, they
865 look like Brownian motion), over long intervals of time sample paths of X_t will begin to appear to have
866 periods of continuity interrupted by an occasional discontinuous jump and thus approach a qualitatively
867 distinct process. These emergent properties can be formally characterized by Lévy processes and have
868 been successfully employed in comparative phylogenetics to fit phenotypic data from extant populations
869 and the fossil record (Landis and Schraiber 2017). It would therefore be interesting to investigate whether
870 an application of a separation of time scales argument for the rate of environmental change (λ) versus the
871 rate of evolutionary and ecological change (a) to microevolutionary models derived using our framework
872 can be used to obtain macroevolutionary models that include not only mean trait evolution, but also the
873 evolution of trait variance and abundance. The resulting macroevolutionary models can give rise to novel
874 comparative phylogenetic methods and provide initial conditions for microevolutionary models that capture
875 contemporary dynamics.

876 Final remarks

877 Although top-down approaches to community ecology, such as the Maximum Entropy Theory of Ecology
878 (Harte 2011), have enjoyed some success in describing community-level patterns (Harte and Newman 2014;
879 Xiao, McGlinn, and White 2015), a mechanistic understanding of why these patterns emerge and how
880 they will change remains a formidable task. Such an understanding must take both bottom-up and top-
881 down approaches integrating considerations from the ecophysiology of individual organisms that reveal the
882 economics of interspecific interactions (Sterner and Elser 2008), to the phylogeographic history of taxa that
883 sets the stage for contemporary dynamics (Hickerson et al. 2010). Through connecting these dots we can
884 increase the variance explained in observations of ecological communities by specific mechanisms and come
885 closer to a predictive theory of evolutionary community ecology. Despite the long list of equations derived
886 in this paper, this work takes just one small step towards capturing these many details. However, we hope
887 the theoretical framework outlined here along with the demonstration of its use in modelling competitive
888 communities provides some helpful tools to aid quantitative evolutionary ecologists in reaching such lofty
889 goals.

890 5 Supplementary material (SM)

891 Throughout this supplement, we set use dot notation for time derivatives so that $\dot{f}(x, t) = \frac{\partial}{\partial t} f(x, t)$ and set
 892 $\Delta = \frac{\partial^2}{\partial x^2}$, except in §5.8.1.3 where Δ represents a random variable.

893 5.1 Sufficient conditions for finite mean, variance and total abundance in the 894 deterministic case

895 Recall $C_{1,c}^+(\mathbb{R} \times [0, \infty))$ denotes the set of non-negative functions $\nu(x, t)$ continuous in both x and t such that
 896 $\int_{\mathbb{R}} \nu(x, t) dx < +\infty$. We assume, for $m : C_{1,c}^+(\mathbb{R} \times [0, \infty)) \times \mathbb{R} \rightarrow \mathbb{R}$, there exists $r \in \mathbb{R}$ such that $m(\nu, x) \leq r$
 897 for each $\nu \in C_{1,c}^+(\mathbb{R} \times [0, \infty))$ and $x \in \mathbb{R}$. As in the main text, we also assume the initial condition $\nu(x, 0)$
 898 satisfies

$$899 N(0) = \int_{\mathbb{R}} \nu(x, 0) dx < +\infty, \quad (55)$$

$$900 -\infty < \bar{x}(0) = \int_{\mathbb{R}} x \nu(x, 0) dx < +\infty, \quad (56)$$

$$901 \sigma^2(0) = \int_{\mathbb{R}} (x - \bar{x}(0))^2 \nu(x, 0) dx < +\infty, \quad (57)$$

902 where $p(x, 0) = \nu(x, 0)/N(0)$, and we consider the PDE

$$903 \dot{\nu}(x, t) = m(\nu, x) \nu(x, t) + \frac{\mu}{2} \Delta \nu(x, t). \quad (58)$$

904 Replacing m with its upper bound $r \in \mathbb{R}$, equation (58) reduces to a simple parabolic equation that can be solved using standard techniques (Farlow 1993). In particular, when $m(\nu, x) \equiv 0$ denote the solution to (58) by $\nu_0(x, t)$. Then, denoting

$$905 \Phi(x, t) = \frac{\exp(-x^2/2\mu t)}{\sqrt{2\pi\mu t}}, \quad (59)$$

906 we have

$$907 \nu_0(x, t) = \int_{\mathbb{R}} \Phi(x - y, t) \nu(y, 0) dy. \quad (60)$$

908 In the more general case, when $m(\nu, x) \equiv r \in \mathbb{R}$, equation (58) has the solution $\nu_r(x, t) = e^{rt} \nu_0(x, t)$. Hence, $\nu_r(x, t) \geq 0$ for all $x \in \mathbb{R}$ and $\int_{\mathbb{R}} \nu_r(x, t) dx = e^{rt} N(0) < +\infty$ for all $t \geq 0$. Furthermore, denoting $N_r(t) = \int_{\mathbb{R}} \nu_r(x, t) dx$, $p_r(x, t) = \nu_r(x, t)/N_r(t)$, $\bar{x}_r(t) = \int_{\mathbb{R}} x p_r(x, t) dx$ and $\sigma_r^2(t) = \int_{\mathbb{R}} (x - \bar{x}_r(t))^2 p_r(x, t) dx$, we have

$$909 \bar{x}_r(t) = \int_{\mathbb{R}} x \int_{\mathbb{R}} \Phi(x - y, t) p(y, 0) dy dx = \int_{\mathbb{R}} y p(y, 0) dy = \bar{x}(0), \quad (61)$$

$$910 \sigma_r^2(t) = \int_{\mathbb{R}} (x - \bar{x}_r(t))^2 \int_{\mathbb{R}} \Phi(x - y, t) p(y, 0) dy dx = \int_{\mathbb{R}} ((y - \bar{x}(0))^2 + \mu t) p(y, 0) dy = \sigma^2(0) + \mu t. \quad (62)$$

911 Hence, $|\bar{x}_r(t)|, \sigma_r^2(t) < +\infty$ for all $t \geq 0$. For the sake of contradiction, suppose there exists $x \in \mathbb{R}$ and $t \geq 0$ such that $\nu(x, t) > \nu_r(x, t)$. Then

$$912 \nu(x, t) - \nu(x, 0) = \int_0^t m(\nu, s) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) ds > \int_0^t r \nu_r(x, s) + \frac{\mu}{2} \Delta \nu_r(x, s) ds = \nu_r(x, t) - \nu(x, 0) \quad (63)$$

913 which implies there exists $\nu \in C_{1,c}^+(\mathbb{R} \times [0, \infty))$ and $x \in \mathbb{R}$ such that $m(\nu, x) > r$. But this contradicts our assumption $m(\nu, x) \leq r$ for all $\nu \in C_{1,c}^+(\mathbb{R} \times [0, \infty))$ and $x \in \mathbb{R}$. So we have $\nu(x, t) \leq \nu_r(x, t)$ for each $x \in \mathbb{R}$ and $t \geq 0$. This implies that $N(t) = \int_{\mathbb{R}} \nu(x, t) dx < +\infty$,

$$914 0 < \int_{\mathbb{R}} x^2 \nu(x, t) dx \leq \int_{\mathbb{R}} x^2 \nu_r(x, t) dx < +\infty \quad (64)$$

⁹¹⁶ and in particular

$$0 < \sigma^2(t) + \bar{x}^2(t) = \frac{1}{N(t)} \int_{\mathbb{R}} x^2 \nu(x, t) dx < +\infty \quad (65)$$

⁹¹⁷ for each $t \geq 0$.

⁹¹⁸ 5.2 The relation between diffusion and convolution with a Gaussian kernel

⁹¹⁹ Let $g : \mathbb{R}^d \rightarrow \mathbb{R}$ be smooth. Consider the deterministic Cauchy problem

$$\begin{cases} \dot{f}(x, t) = -\Delta f(x, t), & (x, t) \in \mathbb{R}^d \times (0, \infty) \\ f(x, t) = g(x), & (x, t) \in \mathbb{R}^d \times \{0\}. \end{cases} \quad (\text{SM1.1})$$

⁹²⁰ According to Evans (2010), the fundamental solution of (SM1.1) is

$$\Phi(x, t) = \frac{1}{(4\pi t)^{d/2}} \exp\left(-\frac{|x|^2}{4t}\right), \quad (x, t) \in (0, \infty) \times \mathbb{R}^d, \quad (\text{SM1.2})$$

⁹²¹ where $|x| = \sqrt{\sum_i x_i^2}$. The solution $f(x, t)$ of PDE (SM1.1) is then given by the convolution

$$f(x, t) = \int_{\mathbb{R}^d} \Phi(x - y, t) g(y) dy, \quad (x, t) \in (0, \infty) \times \mathbb{R}^d. \quad (\text{SM1.3})$$

⁹²² Hence, by the fundamental theorem of calculus,

$$\begin{aligned} f(x, t) + \int_t^{t+1} \dot{f}(x, s) ds &= f(x, t+1) \\ &= \int_{\mathbb{R}^d} \Phi(x - y, t+1) g(y) dy = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \Phi(x - y, 1) \Phi(y - z, t) g(z) dz dy \\ &= \int_{\mathbb{R}^d} \Phi(x - y, 1) f(t, y) dy. \end{aligned} \quad (\text{SM1.4})$$

⁹²³ In particular,

$$f(x, t) + \int_t^{t+1} \Delta f(x, s) ds = \int_{\mathbb{R}^d} \Phi(1, x - y) f(y, t) dy. \quad (\text{SM1.5})$$

⁹²⁴ 5.3 Deterministic dynamics of $\sigma^2(t)$

⁹²⁵ Picking up from the main text §2.1,

$$\begin{aligned} \dot{\sigma}^2(t) &= \frac{d}{dt} \int_{\mathbb{R}} (x - \bar{x}(t))^2 p(x, t) dx \\ &= \int_{\mathbb{R}} 2(x - \bar{x}(t)) \dot{x}(t) + (x - \bar{x}(t))^2 \dot{p}(x, t) dx \\ &= \int_{\mathbb{R}} (x - \bar{x}(t))^2 \left((m(\nu, x) - \bar{m}(t)) p(x, t) + \frac{\mu}{2} \frac{\partial^2}{\partial x^2} p(x, t) \right) dx \end{aligned} \quad (66)$$

$$\begin{aligned}
&= \int_{\mathbb{R}} ((x - \bar{x}(t))^2 - \sigma^2(t) + \sigma^2(t)) (m(\nu, x) - \bar{m}(t)) p(x, t) + (x - \bar{x}(t))^2 \frac{\mu}{2} \frac{\partial^2}{\partial x^2} p(x, t) dx \\
&= \text{Cov}_t((x - \bar{x}(t))^2, m(\nu, x)) + \frac{\mu}{2} \int_{\mathbb{R}} (x - \bar{x}(t))^2 \frac{\partial^2}{\partial x^2} p(x, t) dx.
\end{aligned}$$

⁹²⁶ In particular, when $p(x, t)$ is Gaussian,

$$\frac{\partial^2}{\partial x^2} p(x, t) = \frac{(x - \bar{x}(t))^2 - \sigma^2(t)}{\sigma^4(t)} p(x, t) \quad (67)$$

⁹²⁷ and hence, referring to the well-known moments of the Gaussian distribution, we find

$$\begin{aligned}
&\int_{\mathbb{R}} (x - \bar{x}(t))^2 \frac{\partial^2}{\partial x^2} p(x, t) dx \\
&= \int_{\mathbb{R}} \frac{(x - \bar{x}(t))^4 - (x - \bar{x}(t))^2 \sigma^2(t)}{\sigma^4(t)} p(x, t) dx = 2.
\end{aligned} \quad (68)$$

⁹²⁸ In the Gaussian case, we also have

$$\begin{aligned}
2\sigma^4 \left(\frac{\partial \bar{m}}{\partial \sigma^2} - \overline{\frac{\partial m}{\partial \sigma^2}} \right) &= 2\sigma^4 \left(\frac{\partial}{\partial \sigma^2} \int_{\mathbb{R}} m(\nu, x) p(x, t) dx - \int_{\mathbb{R}} p(x, t) \frac{\partial}{\partial \sigma^2} m(\nu, x) dx \right) \quad (69) \\
&= 2\sigma^4 \int_{\mathbb{R}} \frac{(x - \bar{x})^2 - \sigma^2}{2\sigma^4} m(\nu, x) p(x, t) dx = \int_{\mathbb{R}} ((x - \bar{x})^2 - \sigma^2) (m(\nu, x) - \bar{m}) p(x, t) dx \\
&= \text{Cov}_t((x - \bar{x})^2, m).
\end{aligned}$$

⁹²⁹ 5.4 Numerical evidence of finite moments and approximate normality in the ⁹³⁰ stochastic case

⁹³¹ Here we use a numerical argument to suggest, for

$$r - \frac{a}{2}(\theta - x)^2 - c \int_{\mathbb{R}} \nu(x, t) dx \leq m(\nu, x) \leq r - \frac{a}{2}(\theta - x)^2, \quad \forall (\nu, x) \in C_{1,c}^+(\mathbb{R}) \times \mathbb{R}, \quad (70)$$

⁹³² the density process $\nu(x, t)$ defined by SPDE (25) of the main text satisfies $\int_{\mathbb{R}} (|x| + x^2) \nu(x, t) dx < \infty$. From
⁹³³ SM §5.1, under the assumption $m(\nu, x) = r - \frac{a}{2}(\theta - x)^2 - c \int_{\mathbb{R}} \nu(x, t) dx$, we can derive the differential equations

$$\dot{\bar{x}} = aG(\theta - \bar{x}) \quad (71a)$$

$$\dot{G} = \mu - aG^2 \quad (71b)$$

$$\dot{N} = \left(r - \frac{a}{2} ((\theta - \bar{x})^2 + G + \eta) - cN \right) N. \quad (71c)$$

⁹³⁴ Ignoring the trivial case of $N = 0$, the equilibrium

$$\hat{\dot{x}} = \theta, \quad (72a)$$

$$\hat{G} = \sqrt{\frac{\mu}{a}}, \quad (72b)$$

$$\hat{N} = \frac{1}{c} \left(r - \frac{1}{2}(\eta a + \sqrt{\mu a}) \right), \quad (72c)$$

939 is unique and globally stable for $a, c, \mu > 0$. Setting $2r > \eta a + \sqrt{\mu a}$ ensures a positive equilibrium abundance
940 and setting $c < r - (\eta a + \sqrt{\mu a})/2$ ensures $\hat{N} > 1$, which is important for numerical simulations when N is an
941 integer. We use these results to help guide our choice of parameter values for simulations of the branching
942 random walk. In the following section we provide a detailed description of the branching random walk and
943 how we have chosen to rescale it. We then use the rescaled branching random walk to investigate finiteness
944 of moments and normality.

945 5.4.1 Description of simulation

946 We begin by describing the branching random walk before introducing our scheme to rescale it. Our branching
947 random walk follows closely the description of branching Brownian motion in the main text. However, we
948 replace exponentially distributed lifetimes with deterministic unit time steps for easier implementation.
949 Hence, we restrict time to $t = 0, 1, 2, \dots$. Furthermore, we allow individual fitness to depend on both trait
950 value and the state of the entire population. For time t we write $\{\xi_1(t), \dots, \xi_{N(t)}(t)\}$ as the set of trait
951 values across all $N(t)$ individuals alive in the population. Since our simulation treats individuals instead of
952 continuous distributions of trait values, we can write

$$\nu(x, t) = \sum_{i=1}^{N(t)} \delta(x - \xi_i(t)), \quad (73)$$

953 where $\delta(x - \xi_i(t))$ denotes the point mass located at $\xi_i(t)$. To allow for imperfect heritability, we also track
954 the set of breeding values which, at time t , is denoted by $\{\gamma_1(t), \dots, \gamma_{N(t)}(t)\}$ and should not be confused
955 with the quadratic selection gradients discussed in §?? of the main text. Then the distribution of breeding
956 values can be written as

$$\rho(g, t) = \sum_{i=1}^{N(t)} \delta(g - \gamma_i(t)). \quad (74)$$

957 Following our model of heritability, the trait value $\xi_i(t)$ is drawn from a normal distribution centered on
958 $\gamma_i(t)$ with variance η . At each iteration we draw, for each individual, a random number of offspring from
959 a Negative-Binomial distribution. Recall the Negative-Binomial distribution models the number of failed
960 Bernoulli trials that occur before a given number of successful trials. Denoting q the probability of success
961 for each trial and s the number of successes, the mean and variance is given respectively by

$$\frac{s(1-q)}{q}, \frac{s(1-q)}{q^2}. \quad (75)$$

962 Then if we require the i th individual to have mean number offspring $\mathcal{W}(\nu, \xi_i)$ and variance equal to V , the
963 parameters of the associated Negative-Binomial distribution become

$$q(\nu, \xi_i) = \frac{\mathcal{W}(\nu, \xi_i)}{V}, \quad s(\nu, \xi_i) = \frac{\mathcal{W}^2(\nu, \xi_i)}{V - \mathcal{W}(\nu, \xi_i)}. \quad (76)$$

964 The imposes the restriction $V > \mathcal{W}(\nu, \xi_i)$. For each offspring produced by the individual with breeding value
965 $\gamma_i(t)$, we assign independently drawn breeding values normally distributed around $\gamma_i(t)$ with variance μ . After

breeding values have been assigned, we randomly draw trait values for each offspring as described above.
 For an overview of our model of inheritance, see §?? of the main text. This summarizes the basic structure
 of our simulation. To impose selection and density dependent growth rates, we set

$$\mathcal{W}(\nu, \xi_i) = \exp \left(r - \frac{a}{2} (\theta - \xi_i)^2 - c \int_{\mathbb{R}} \nu(x, t) dx \right), \quad (77)$$

where the above integral becomes $\int_{\mathbb{R}} \nu(x, t) dx = \sum_{i=1}^{N(t)} 1 = N(t)$.

Rescaling

To rescale the branching random walk by a positive integer n , we rescale segregation and mutational variance according to $\eta \rightarrow \eta$ and $\mu \rightarrow \mu/n$, time by $t \rightarrow t/n$ and the reproductive law by $V \rightarrow V$ and

$$\mathcal{W}(\nu, \xi_i) \rightarrow \mathcal{W}^{(n)}(\nu, \xi_i) = \exp \left(\frac{r}{n} - \frac{a}{2n} (\theta - \xi_i)^2 - \frac{c}{n^2} N(t) \right) = \exp \left(\frac{r}{n} - \frac{a}{2n} (\theta - \xi_i)^2 - \frac{c}{n} N^{(n)}(t) \right). \quad (78)$$

We also replace individual mass with $\frac{1}{n}$ and write rescaled abundance as $N^{(n)}(t) = \frac{1}{n} N(nt)$. Under this rescaling the deterministic equilibrium of the raw numerical abundance becomes

$$\hat{N} = \frac{n^2}{c} \left(\frac{r}{n} - \frac{1}{2n} (\eta a + \sqrt{\mu a}) \right) = \frac{n}{c} \left(r - \frac{1}{2} (\eta a + \sqrt{\mu a}) \right). \quad (79)$$

The deterministic equilibrium of the rescaled abundance is then

$$\hat{N}^{(n)} = \frac{1}{c} \left(r - \frac{1}{2} (\eta a + \sqrt{\mu a}) \right). \quad (80)$$

When it exists, we denote by $N^{(\infty)}(t)$ the limiting process of $N^{(n)}(t)$. Then

$$\lim_{n \rightarrow \infty} n \left(\mathcal{W}^{(n)}(\nu, \xi_i) - 1 \right) = r - \frac{a}{2} (\theta - \xi_i)^2 - c N^{(\infty)}(t). \quad (81)$$

Note that, following the notation of Theorem 1 in Méléard and Roelly (1992), setting $\lambda_n = n$, $m_n(\nu) = \mathcal{W}^{(n)}(\nu, \cdot)$ and $\varepsilon_n = 1/n$ satisfies their hypotheses (\mathcal{H}_0) - (\mathcal{H}_3) when $c = 0$. We have implemented this simulation in the programming language Julia. A copy can be found at the url:

<https://github.com/bobweek/branching.brownian.motion.and.spde>

For the sake of illustration, we simulated the unscaled process ($n = 1$) and the rescaled process with $n = 5$ and $n = 20$ for 50 units of time. Results are shown in Figure 2. In the following section we use a statistical test to show, for the lower bound on $m(\nu, x)$, the rescaled process converges to a Gaussian density as $n \rightarrow \infty$ and $V/N \rightarrow 0$.

5.4.2 Evidence of normality

To demonstrate approximate normality of the phenotypic distribution when V/N is small we utilized the one-sided Kolmogorov-Smirnov test. This test compares an empirical cumulative distribution function (i.e., a cumulative distribution function generated from simulated data) to a hypothetical cumulative distribution function by providing a distribution for the maximum distance between these curves. More precisely if $F_n(x)$ is the empirical distribution function for a sample of size n and $F(x)$ is the hypothetical distribution function, Kolmogorov-Smirnov statistic is $D_n = \sup_x |F_n(x) - F(x)|$.

⁹⁹² **5.5 Derivation of SDE for \bar{x} and σ^2**

⁹⁹³ For $\nu(x, t)$ defined in the main text, $h \in C(\mathbb{R})$ and $t \geq 0$ we make the following assumptions:

$$\mathbb{E} \left(\int_{\mathbb{R}} |h(x)| \nu(x, t) dx \right) < \infty, \quad (82)$$

$$\mathbb{E} \left(\int_{\mathbb{R}} h^2(x) \nu(x, t) dx \right) < \infty, \quad (83)$$

$$\mathbb{E} \left(\int_0^t \int_{\mathbb{R}} \nu(x, s) \left| h(x) \right| \left| m(\nu, x) f(x) + \frac{\mu}{2} \Delta f(x) \right| dx ds \right) < \infty, \quad \forall f \in C_b^2(\mathbb{R}). \quad (84)$$

⁹⁹⁶ Put $H(t) = \int_{\mathbb{R}} h(x) \nu(x, t) dx$. Then, for non-random and non-negative $\nu_1(x, t)$ that is continuous in both arguments and integrable in x ,

$$\star := \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E}[H(t + \varepsilon) - H(t) | \nu(x, t) = \nu_1(x, t)] \quad (85)$$

$$= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E} \left[\int_{\mathbb{R}} h(x) (\nu(x, t + \varepsilon) - \nu(x, t)) dx \middle| \nu(x, t) = \nu_1(x, t) \right] \quad (86)$$

$$= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E} \left[\int_{\mathbb{R}} \int_t^{t+\varepsilon} f(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) + \sqrt{V \nu(x, s)} \dot{W}(x, s) \right) ds dx \middle| \nu(x, t) = \nu_1(x, t) \right]. \quad (87)$$

¹⁰⁰⁰ By assumption (84) we can use Fubini's theorem to write, with probability one,

$$\int_{\mathbb{R}} \int_t^{t+\varepsilon} h(x) (m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s)) ds dx = \int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) (m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s)) dx ds. \quad (88)$$

¹⁰⁰¹ By assumption (83) we have $\varphi(x, t) = h(x) \sqrt{V \nu(x, t)}$ implies $\varphi \in L_c^2(\mathbb{R} \times [0, \infty))$. Hence, the following is true by definition;

$$\int_{\mathbb{R}} \int_t^{t+\varepsilon} h(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) ds dx = \int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) dx ds. \quad (89)$$

¹⁰⁰³ Hence,

$$\begin{aligned} \mathbb{E} \left[\int_{\mathbb{R}} \int_t^{t+\varepsilon} h(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) ds dx \middle| \nu(x, t) = \nu_1(x, t) \right] \\ = \mathbb{E} \left[\int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) dx ds \middle| \nu(x, t) = \nu_1(x, t) \right] = 0 \end{aligned} \quad (90)$$

¹⁰⁰⁴ and

$$\star = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E} \left[\int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) \right) dx ds \middle| \nu(x, t) = \nu_1(x, t) \right]. \quad (91)$$

¹⁰⁰⁵ By assumption (84) we know that there exists a $\delta > 0$ such that for each positive $\varepsilon < \delta$ the following holds almost surely:

$$\begin{aligned} & \left| \int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) \right) dx ds \right| \\ & \leq \int_t^{t+\delta} \int_{\mathbb{R}} |h(x)| \left| m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) \right| dx ds < \infty. \end{aligned} \quad (92)$$

1007 Thus, by Lebesgue's dominated convergence theorem, the drift component of the process $H(t)$ can be com-
1008 puted as

$$\begin{aligned} \star = \mathbb{E} \left[\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) \right) dx ds \middle| \nu(x, t) = \nu_1(x, t) \right] \\ = \int_{\mathbb{R}} h(x) \left(m(\nu_1, x) \nu_1(x, t) + \frac{\mu}{2} \Delta \nu_1(x, t) \right) dx. \end{aligned} \quad (93)$$

1009 To find an expression for the diffusion component of $H(t)$ set

$$\begin{aligned} \star \star := \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{V}[H(t + \varepsilon) - H(t) | \nu(x, t) = \nu_1(x, t)] \\ = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{V} \left[\int_{\mathbb{R}} h(x) (\nu(x, t + \varepsilon) - \nu(x, t)) dx \middle| \nu(x, t) = \nu_1(x, t) \right]. \end{aligned} \quad (94)$$

1010 We can rewrite the integral inside expression (94) as

$$\int_{\mathbb{R}} \int_t^{t+\varepsilon} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) + \sqrt{V \nu(x, s)} \dot{W}(x, s) \right) ds dx. \quad (95)$$

1011 We have already found

$$\begin{aligned} & \mathbb{E} \left[\int_{\mathbb{R}} \int_t^{t+\varepsilon} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) + \sqrt{V \nu(x, s)} \dot{W}(x, s) \right) ds dx \middle| \nu(x, t) = \nu_1(x, t) \right] \\ & = \mathbb{E} \left[\int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) \right) dx ds \middle| \nu(x, t) = \nu_1(x, t) \right]. \end{aligned} \quad (96)$$

1012 Then, since $h^2(x) \nu(x, t)$ is integrable by assumption (83), we have

$$\begin{aligned} \star \star = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E} \left[\left(\int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) ds dx \right)^2 \middle| \nu(x, t) = \nu_1(x, t) \right] \\ = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E} \left[\int_t^{t+\varepsilon} \int_{\mathbb{R}} V h^2(x) \nu(x, s) dx ds \middle| \nu(x, t) = \nu_1(x, t) \right]. \end{aligned} \quad (97)$$

1013 Thus, for any $\delta \geq \varepsilon \geq 0$, we have, with probability one,

$$\int_t^{t+\varepsilon} \int_{\mathbb{R}} V h^2(x) \nu(x, s) dx ds \leq \int_t^{t+\delta} \int_{\mathbb{R}} V h^2(x) \nu(x, s) dx ds. \quad (98)$$

1014 We can therefore use Lebesgue's dominated convergence theorem to justify

$$\star\star = \mathbb{E} \left[\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_t^{t+\varepsilon} \int_{\mathbb{R}} V h^2(x) \nu(x, s) dx ds \middle| \nu(x, t) = \nu_1(x, t) \right] = \int_{\mathbb{R}} V h^2(x) \nu_1(x, t) dx. \quad (99)$$

1015 Then, using the notation of stochastic differentials, we have

$$dH(t) = \left(\int_{\mathbb{R}} h(x) \left(m(\nu, x) \nu(x, t) + \frac{\mu}{2} \Delta \nu(x, t) \right) dx \right) dt + \sqrt{V \int_{\mathbb{R}} h^2(x) \nu(x, t) dx} dW(t) \quad (100)$$

1016 where W is a standard Brownian motion. In the following subsections we employ this formula under the
1017 cases $h(x) = x, x^2$ to obtain SDE for the phenotypic mean and variance.

1018 5.5.1 Derivation for trait mean

1019 We set $\tilde{x}(t) = \int_{\mathbb{R}} x \nu(x, t)$ and make use of the notation

$$\begin{aligned} \|N\|_2 &= \sqrt{V \int_{\mathbb{R}} \nu(x, t) dx} = \sqrt{VN} \\ \|\tilde{x}\|_2 &= \sqrt{V \int_{\mathbb{R}} x^2 \nu(x, t) dx} \\ \langle \tilde{x}, N \rangle &= V \int_{\mathbb{R}} x \nu(x, t) dx = \bar{x} VN. \end{aligned} \quad (101)$$

1020 Applying formula (100) provides

$$d\tilde{x} = \left(\bar{x}mN + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta \nu(x, t) dx \right) dt + \|\tilde{x}\|_2 d\tilde{W}_2, \quad (102)$$

1021 where,

$$d\tilde{W}_2 = d\hat{\mathbf{W}}_{\sqrt{Vx^2\nu}} = \frac{1}{\|\tilde{x}\|_2} \int_{\mathbb{R}} x \sqrt{V\nu(x, t)} \dot{W}(x, t) dx dt. \quad (103)$$

1022 Using Itô's quotient rule on $\bar{x} = \tilde{x}/N$, we obtain

$$d\bar{x} = d\left(\frac{\tilde{x}}{N}\right) = \frac{\tilde{x}}{N} \left(\frac{d\tilde{x}}{\tilde{x}} - \frac{dN}{N} - \frac{d\tilde{x}}{\tilde{x}} \frac{dN}{N} + \left(\frac{dN}{N}\right)^2 \right) = \frac{d\tilde{x}}{N} - \bar{x} \frac{dN}{N} - \frac{d\tilde{x}}{N} \frac{dN}{N} + \bar{x} \left(\frac{dN}{N}\right)^2. \quad (104)$$

1023 From Table 1 of the main text $d\tilde{x}dN = \langle \tilde{x}, N \rangle$ and $dN^2 = \|N\|_2^2$. Hence,

$$\begin{aligned} d\bar{x} &= \bar{x}mdt + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x, t) dx dt + \frac{\|\tilde{x}\|_2}{N} d\tilde{W}_2 - \bar{x} \left(\bar{m}dt + \sqrt{\frac{V}{N}} dW_1 \right) - \frac{\langle \tilde{x}, N \rangle}{N^2} dt + \bar{x} \frac{\|N\|_2^2}{N^2} dt \\ &= (\bar{x}m - \bar{x}\bar{m})dt + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x, t) dx dt + \frac{\|\tilde{x}\|_2}{N} d\tilde{W}_2 - \bar{x} \sqrt{\frac{V}{N}} dW_1 - V \frac{\bar{x}}{N} dt + V \frac{\bar{x}}{N} dt \\ &= \left(\text{Cov}_t(x, m) + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x, t) dx \right) dt + \frac{\|\tilde{x}\|_2}{N} d\tilde{W}_2 - \bar{x} \sqrt{\frac{V}{N}} dW_1. \end{aligned} \quad (105)$$

1024 Note that

$$\begin{aligned} \frac{\|\tilde{x}\|_2}{N} d\tilde{W}_2 - \bar{x} \sqrt{\frac{V}{N}} dW_1 &= \frac{1}{N} \int_{\mathbb{R}} x \sqrt{V\nu(x,t)} \dot{W}(x,t) dx - \frac{\bar{x}}{N} \int_{\mathbb{R}} \sqrt{V\nu(x,t)} \dot{W}(x,t) dx \\ &= \int_{\mathbb{R}} \frac{x - \bar{x}}{N} \sqrt{V\nu(x,t)} \dot{W}(x,t) dx \end{aligned} \quad (106)$$

1025 and

$$\mathbb{V} \left(\int_{\mathbb{R}} \frac{x - \bar{x}}{N} \sqrt{V\nu(x,t)} \dot{W}(x,t) dx \right) = \frac{V}{N} \int_{\mathbb{R}} (x - \bar{x})^2 p(x,t) dx = V \frac{\sigma^2}{N}. \quad (107)$$

1026 Hence, by setting

$$dW_2 = \frac{\int_{\mathbb{R}} \frac{(x - \bar{x})}{N} \sqrt{V\nu(x,t)} \dot{W}(x,t) dx}{\sqrt{V\sigma^2/N}} \quad (108)$$

1027 we can write

$$d\bar{x} = \left(\text{Cov}_t(x, m) + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x,t) dx \right) dt + \sqrt{V \frac{\sigma^2}{N}} dW_2. \quad (109)$$

1028 5.5.2 Derivation for trait variance

1029 We set $\tilde{\sigma}^2(t) = \int_{\mathbb{R}} x^2 \nu(x,t) dx$ and make use of the notation

$$\begin{aligned} \|\tilde{\sigma}^2\|_2 &= \sqrt{V \int_{\mathbb{R}} x^4 \nu(x,t) dx} \\ \langle \tilde{\sigma}^2, N \rangle &= V \int_{\mathbb{R}} x^2 \nu(x,t) dx = \bar{x}^2 VN. \end{aligned} \quad (110)$$

1030 Applying formula (100) provides

$$d\tilde{\sigma}^2 = \left(\bar{x}^2 m N + \frac{\mu}{2} \int_{\mathbb{R}} x^2 \Delta \nu(x,t) dx \right) dt + \|\tilde{\sigma}^2\|_2 d\tilde{W}_3 \quad (111)$$

1031 where

$$d\tilde{W}_3 = d\hat{\mathbf{W}}_{\sqrt{Vx^4\nu}} = \frac{1}{\|\tilde{\sigma}^2\|_2} \int_{\mathbb{R}} x^2 \sqrt{V\nu(x,t)} \dot{W}(x,t) dx. \quad (112)$$

1032 Using Itô's quotient rule on $\bar{x}^2 = \tilde{\sigma}^2/N$, we obtain

$$d\bar{x}^2 = d\left(\frac{\tilde{\sigma}^2}{N}\right) = \frac{\tilde{\sigma}^2}{N} \left(\frac{d\tilde{\sigma}^2}{\tilde{\sigma}^2} - \frac{dN}{N} - \frac{d\tilde{\sigma}^2}{\tilde{\sigma}^2} \frac{dN}{N} + \left(\frac{dN}{N}\right)^2 \right) = \frac{d\tilde{\sigma}^2}{N} - \frac{\bar{x}^2}{N} \frac{dN}{N} - \frac{d\tilde{\sigma}^2}{N} \frac{dN}{N} + \frac{\bar{x}^2}{N} \left(\frac{dN}{N}\right)^2. \quad (113)$$

1033 Table 1 of the main text implies $d\tilde{W}_3 dW_1 = \langle \tilde{\sigma}^2, N \rangle$ and hence

$$\begin{aligned}
d\bar{x}^2 &= \left(\bar{x}^2 \bar{m} + \frac{\mu}{2} \int_{\mathbb{R}} x^2 \Delta p(x, t) dx \right) dt + \frac{\|\tilde{\sigma}^2\|_2}{N} d\tilde{W}_3 - \bar{x}^2 \left(\bar{m} dt + \sqrt{\frac{V}{N}} dW_1 \right) - \frac{\langle \tilde{\sigma}^2, N \rangle}{N^2} dt + \bar{x}^2 \frac{\|N\|_2^2}{N^2} dt \\
&= \left(\bar{x}^2 \bar{m} - \bar{x}^2 \bar{m} dt + \frac{\mu}{2} \int_{\mathbb{R}} x^2 \Delta p(x, t) dx \right) dt + \frac{\|\tilde{\sigma}^2\|_2}{N} d\tilde{W}_3 - \bar{x}^2 \sqrt{\frac{V}{N}} dW_1 - \bar{x}^2 \frac{V}{N} dt + \bar{x}^2 \frac{V}{N} dt \\
&= \left(\text{Cov}_t(x^2, m) + \frac{\mu}{2} \int_{\mathbb{R}} x^2 \Delta p(x, t) dx \right) dt + \frac{\|\tilde{\sigma}^2\|_2}{N} d\tilde{W}_3 - \bar{x}^2 \sqrt{\frac{V}{N}} dW_1. \quad (114)
\end{aligned}$$

¹⁰³⁴ Setting $F(y, z) = y - z^2$, use Itô's formula on $\sigma^2 = F(\bar{x}^2, \bar{x}) = \bar{x}^2 - \bar{x}^2$ to obtain:

$$\begin{aligned}
d\sigma^2 &= d\bar{x}^2 - 2\bar{x}d\bar{x} - (d\bar{x})^2 = \left(\text{Cov}_t(x^2, m) + \frac{\mu}{2} \int_{\mathbb{R}} x^2 \Delta p(x, t) dx \right) dt + \frac{\|\tilde{\sigma}^2\|_2}{N} d\tilde{W}_3 - \bar{x}^2 \sqrt{\frac{V}{N}} dW_1 \\
&- 2\bar{x} \left(\text{Cov}_t(x, m) dt + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x, t) dx dt + \sqrt{\frac{V\sigma^2}{N}} dW_2 \right) - \left(\text{Cov}_t(x, m) dt + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x, t) dx dt + \sqrt{\frac{V\sigma^2}{N}} dW_2 \right)^2 \\
&= \left(\text{Cov}_t(x^2 - 2\bar{x}x, m) + \frac{\mu}{2} \int_{\mathbb{R}} (x^2 - x\bar{x}) \Delta p(x, t) dx \right) dt + \frac{\|\tilde{\sigma}^2\|_2}{N} d\tilde{W}_3 - \bar{x}^2 \sqrt{\frac{V}{N}} dW_1 - 2\bar{x} \sqrt{\frac{V\sigma^2}{N}} dW_2 - \left(\frac{V\sigma^2}{N} \right) dt \\
&= \left(\text{Cov}_t((x - \bar{x})^2, m) + \frac{\mu}{2} \int_{\mathbb{R}} (x - \bar{x})^2 \Delta p(x, t) dx - \frac{V\sigma^2}{N} \right) dt + \frac{\|\tilde{\sigma}^2\|_2}{N} d\tilde{W}_3 - \bar{x}^2 \sqrt{\frac{V}{N}} dW_1 - 2\bar{x} \sqrt{\frac{V\sigma^2}{N}} dW_2. \quad (115)
\end{aligned}$$

¹⁰³⁵ In light of

$$\begin{aligned}
\frac{\|\tilde{\sigma}^2\|_2}{N} d\tilde{W}_3 - \bar{x}^2 \sqrt{\frac{V}{N}} dW_1 - 2\bar{x} \sqrt{\frac{V\sigma^2}{N}} dW_2 &= \frac{1}{N} \int_{\mathbb{R}} (x^2 - \sigma^2 - 2\bar{x}(x - \bar{x})) \sqrt{V\nu(x, t)} \dot{W}(x, t) dx \\
&= \frac{1}{N} \int_{\mathbb{R}} ((x - \bar{x})^2 - \sigma^2) \sqrt{V\nu(x, t)} \dot{W}(x, t) dx \quad (116)
\end{aligned}$$

¹⁰³⁶ and

$$\begin{aligned}
\frac{1}{N} \int_{\mathbb{R}} \left(((x - \bar{x})^2 - \sigma^2) \sqrt{V\nu(x, s)} \right)^2 dx &= \frac{V}{N} \left(\int_{\mathbb{R}} ((x - \bar{x})^4 - 2(x - \bar{x})^2 \sigma^2 + \sigma^4) p(x, t) dx \right) \\
&= \frac{V}{N} \left(\overline{(x - \bar{x})^4} - \sigma^4 \right) \quad (117)
\end{aligned}$$

¹⁰³⁷ we set

$$dW_3 = \frac{\int_{\mathbb{R}} ((x - \bar{x})^2 - \sigma^2) \sqrt{V\nu(x, t)} \dot{W}(x, t) dx}{V \left(\overline{(x - \bar{x})^4} - \sigma^4 \right)} \quad (118)$$

1038 so that

$$d\sigma^2 = \left(\text{Cov}_t((x - \bar{x})^2, m) + \frac{\mu}{2} \int_{\mathbb{R}} (x - \bar{x})^2 \Delta p(x, t) dx - V \frac{\sigma^2}{N} \right) dt + \sqrt{V \frac{(x - \bar{x})^4 - \sigma^4}{N}} dW_3. \quad (119)$$

1039 Table 1 of the main text implies

$$dW_1 dW_2 = \frac{\int_{\mathbb{R}} (x - \bar{x}) \nu(x, t) dx}{\sqrt{N \sigma^2}} dt = 0, \quad (120)$$

$$dW_1 dW_3 = \frac{\int_{\mathbb{R}} ((x - \bar{x})^2 - \sigma^2) \nu(x, t) dx}{\sqrt{(x - \bar{x})^4 - \sigma^4}} dt = 0, \quad (121)$$

$$dW_2 dW_3 = \frac{\int_{\mathbb{R}} (x - \bar{x}) ((x - \bar{x})^2 - \sigma^2) p(x, t) dx}{\sqrt{\sigma^2 ((x - \bar{x})^4 - \sigma^4)}} dt = \frac{N \overline{(x - \bar{x})^3}}{\sqrt{\sigma^2 ((x - \bar{x})^4 - \sigma^4)}} dt. \quad (122)$$

1040 In particular, when p is a Gaussian curve $dW_2 dW_3 = 0$.

1041 **5.6 Relating fitness of expressed traits to fitness of breeding values**

$$m^*(\rho, g) = \int_{\mathbb{R}} m(\nu, x) \psi(x, g) dx$$

$$\overline{\frac{\partial m^*}{\partial \bar{x}}} = \int_{\mathbb{R}} \frac{\rho(g, t)}{N(t)} \frac{\partial}{\partial \bar{x}} \int_{\mathbb{R}} m(\nu, x) \psi(x, g) dx dg = \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\rho(g, t)}{N(t)} \psi(x, g) dg \frac{\partial}{\partial \bar{x}} m(\nu, x) dx = \int_{\mathbb{R}} p(x, t) \frac{\partial}{\partial \bar{x}} m(\nu, x) dx = \overline{\frac{\partial m}{\partial \bar{x}}}$$

$$\overline{\frac{\partial m^*}{\partial G}} = \int_{\mathbb{R}} \frac{\rho(g, t)}{N(t)} \frac{\partial}{\partial G} \int_{\mathbb{R}} m(\nu, x) \psi(x, g) dx dg = \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\rho(g, t)}{N(t)} \psi(x, g) dg \frac{\partial m}{\partial G} dx = \int_{\mathbb{R}} p(x, t) \frac{\partial m}{\partial \sigma^2} \frac{\partial \sigma^2}{\partial G} dx = \overline{\frac{\partial m}{\partial \sigma^2}}$$

1042 **5.7 Derivation of diffuse coevolution model**

1043 In this section we provide a derivation of our model of diffuse coevolution driven by competition. Since most
1044 of the work in this derivation has already been completed in Supplementary Material §5.5, we focus here
1045 on deriving the Malthusian fitness m as a function of trait value x . We begin with discrete populations of
1046 individuals. In particular, we begin by assuming population size n_i is an integer for each species $i = 1, \dots, S$
1047 before passing to the large population size limit.

1048 The reduction in fitness for an individual of species i caused by competition is captured multiplicatively
1049 by $0 < C_i \leq 1$. Biologically this assumes all competitors affect individuals of a given species equally by
1050 consuming the same amount of resources. This is a mean-field interaction since any individual that consumes
1051 resources has an effect on the fitness of all other individuals competing for the same resources. Denote by x_{ij}
1052 the trait value of the j -th individual belonging to species i . The set of trait values across all individuals in the
1053 community at time $t \geq 0$ is written $X = \{x_{ij}\}$. We denote by \mathcal{B}_{ij} a function that maps X to the cumulative
1054 effect of all competitive interactions on the fitness of the j -th individual in species i . Since individuals do
1055 not compete with themselves the net multiplicative effects on fitness of both interspecific and intraspecific
1056 competition on the j -th individual in species i can be summarized by

$$\mathcal{B}_{ij}(X) = C_i^{\sum_{l \neq j} \mathcal{O}_{ii}(x_{ij}, x_{il}) + \sum_{k \neq i} \sum_{l=1}^{n_k} \mathcal{O}_{ik}(x_{ij}, x_{kl})}, \quad (123)$$

where \mathcal{O}_{ij} , defined in the main text, measures the overlap in resource use between individuals of species i and j as a function of their niche-centers. Writing $\mathcal{W}_{ij}(X)$ as the average number of offspring left by the j -th individual of species i , we have

$$\mathcal{W}_{ij}(X) = \mathcal{A}_i(x_{ij}) \mathcal{B}_{ij}(X), \quad (124)$$

where $\mathcal{A}_i(x) = \int_{\mathbb{R}} e_i(\zeta) u_i(\zeta, x) d\zeta$ accounts for abiotic selection and e_i has been defined in the main text.

We now turn to a diffusion limit. Since we have more than one population, we take the diffusion limit for each population one at a time starting with population 1. We write $\mathbf{n} = (n_1, \dots, n_S)$. Following Méléard and Roelly (1993, 1992) we rescale generation time and individual mass to $\frac{1}{n_1}$ and mean of the reproductive law to

$$\mathcal{W}_{1j}^{(\mathbf{n})}(X) = \mathcal{A}_1(x_{1j})^{1/n_1} \exp \left(\frac{\ln C_1}{n_1^2} \sum_{l \neq j} \mathcal{O}_{11}(x_{1j}, x_{1l}) + \frac{\ln C_1}{n_1} \sum_{k \neq 1} \frac{1}{n_k} \sum_{l=1}^{n_k} \mathcal{O}_{1k}(x_{1j}, x_{kl}) \right). \quad (125)$$

For large n_1 , we have the approximation

$$\mathcal{W}_{1j}^{(\mathbf{n})}(X) \approx \mathcal{A}_1(x_{1j})^{1/n_1} \left(1 + \frac{\ln C_1}{n_1^2} \sum_{l \neq j} \mathcal{O}_{11}(x_{1j}, x_{1l}) + \frac{\ln C_1}{n_1} \sum_{k \neq 1} \frac{1}{n_k} \sum_{l=1}^{n_k} \mathcal{O}_{1k}(x_{1j}, x_{kl}) \right). \quad (126)$$

Hence

$$\lim_{n_1 \rightarrow \infty} n_1 (\mathcal{W}_{1j}^{(\mathbf{n})}(X) - 1) = \ln \mathcal{A}_1(x_{1j}) + \left(\int_{\mathbb{R}} \mathcal{O}_{11}(x_{1j}, y) \nu_1(y, t) dy + \sum_{k \neq 1} \frac{1}{n_k} \sum_{l=1}^{n_k} \mathcal{O}_{1k}(x_{1j}, x_{kl}) \right) \ln C_1. \quad (127)$$

We write $\lim_{\mathbf{n} \rightarrow \infty}$ for the iterated limit $\lim_{n_S \rightarrow \infty} \dots \lim_{n_1 \rightarrow \infty}$ and, assuming $\nu_i(\cdot, t) \in C_1^+(\mathbb{R})$ for $i = 1, \dots, S$ and $t \in [0, \infty)$, we set $\boldsymbol{\nu} = (\nu_1, \dots, \nu_S)$. Then, for any $\boldsymbol{\nu}$, the the diffusion limits for the remaining populations provides the Malthusian parameter for individuals in species i with trait value x_{1j} as

$$m_1(\boldsymbol{\nu}, x_{1j}) := \lim_{\mathbf{n} \rightarrow \infty} n_1 (\mathcal{W}_{1j}^{(\mathbf{n})}(X) - 1) = \ln \mathcal{A}_1(x) + \left(\sum_{k=1}^S \int_{\mathbb{R}} \mathcal{O}_{1k}(x_{1j}, y) \nu_k(y, t) dy \right) \ln C_1. \quad (128)$$

We compute the average niche overlap of an individual in species i with nich location x across all individuals in species j as

$$\bar{\mathcal{O}}_{ij}(x, t) = \frac{\int_{\mathbb{R}} \mathcal{O}_{ij}(x, y) \nu_j(y, t) dy}{\int_{\mathbb{R}} \nu_j(y, t) dy}. \quad (129)$$

We now assume the resource utilization curves $u_i(\zeta)$ and phenotypic densities $\nu_i(x, t)$ are Gaussian curves for $i = 1, \dots, S$. In this case $\bar{\mathcal{O}}_{ij}(x, t)$ simplifies to

$$\bar{\mathcal{O}}_{ij}(x, t) = \frac{\int_{\mathbb{R}} \mathcal{O}_{ij}(x, y) \nu_j(y, t) dy}{\int_{\mathbb{R}} \nu_j(y, t) dy} = \frac{U_i U_j}{\sqrt{2\pi(w_i + w_j + \sigma_j^2(t))}} \exp \left(-\frac{(x - \bar{x}_j(t))^2}{2(w_i + w_j + \sigma_j^2(t))} \right). \quad (130)$$

¹⁰⁷⁴ Setting

$$\sigma_i^2(t) = G_i(t) + \eta_i, \quad (131a)$$

$$R_i = \ln \left(\frac{Q_i U_i}{\sqrt{1 + A_i w_i}} \right), \quad (131b)$$

$$a_i = \frac{A_i}{1 + A_i w_i}, \quad (131c)$$

$$\tilde{b}_{ij}(t) = \frac{1}{w_i + w_j + \sigma_j^2(t)}, \quad (131d)$$

$$c_i = -\ln C_i, \quad (131e)$$

¹⁰⁷⁵ we get

$$m_i(\boldsymbol{\nu}, x) = R_i - \frac{a_i}{2} (x - \theta_i)^2 - c_i \sum_{j=1}^S N_j(t) U_i U_j \sqrt{\frac{\tilde{b}_{ij}(t)}{2\pi}} e^{-\frac{\tilde{b}_{ij}(t)}{2}(x - \bar{x}_j(t))^2}. \quad (132)$$

¹⁰⁷⁶ Hence, our fitness function satisfies condition (??) of the main text.

¹⁰⁷⁷ For the remainder of the derivation we suppress notation indicating dependency on $\boldsymbol{\nu}$, x and t . From (132)
¹⁰⁷⁸ we calculate

$$\frac{\partial m_i}{\partial \bar{x}_i} = c_i N_i U_i^2 \tilde{b}_{ii} (x - \bar{x}_i) \sqrt{\frac{\tilde{b}_{ii}}{2\pi}} e^{-\frac{\tilde{b}_{ii}}{2}(x - \bar{x}_i)^2} \quad (133)$$

$$\begin{aligned} \frac{\partial m_i}{\partial G_i} &= \frac{c_i N_i U_i^2}{2} \left(\frac{(x - \bar{x}_i)^2 - G_i - \eta_i - 2w_i}{(G_i + \eta_i + 2w_i)^2} \right) \sqrt{\frac{\tilde{b}_{ii}}{2\pi}} e^{-\frac{\tilde{b}_{ii}}{2}(x - \bar{x}_i)^2} \\ &= \frac{c_i N_i U_i^2 \tilde{b}_{ii}^2}{2} ((x - \bar{x}_i)^2 - \sigma_i^2 - 2w_i) \sqrt{\frac{\tilde{b}_{ii}}{2\pi}} e^{-\frac{\tilde{b}_{ii}}{2}(x - \bar{x}_i)^2}. \end{aligned} \quad (134)$$

¹⁰⁷⁹ Note that

$$\begin{aligned} &\sqrt{\frac{\tilde{b}_{ii}}{2\pi}} \exp \left(-\frac{\tilde{b}_{ii}}{2}(x - \bar{x}_i)^2 \right) \sqrt{\frac{1}{2\pi\sigma_i^2}} \exp \left(-\frac{(x - \bar{x}_i)^2}{2\sigma_i^2} \right) \\ &= \sqrt{\frac{1}{2\pi(\sigma_i^2 + 1/\tilde{b}_{ii})}} \sqrt{\frac{\sigma_i^2 + 1/\tilde{b}_{ii}}{2\pi\sigma_i^2/\tilde{b}_{ii}}} \exp \left(-\frac{\sigma_i^2 + 1/\tilde{b}_{ii}}{2\sigma_i^2/\tilde{b}_{ii}} (x - \bar{x}_i)^2 \right) \\ &= \sqrt{\frac{1}{4\pi(\sigma_i^2 + w_i)}} \sqrt{\frac{2(\sigma_i^2 + w_i)}{2\pi\sigma_i^2(\sigma_i^2 + 2w_i)}} \exp \left(-\frac{\sigma_i^2(\sigma_i^2 + 2w_i)}{4(\sigma_i^2 + w_i)} (x - \bar{x}_i)^2 \right). \end{aligned} \quad (135)$$

¹⁰⁸⁰ Hence,

$$\overline{\frac{\partial m_i}{\partial \bar{x}_i}} = 0, \quad (136)$$

$$\begin{aligned}\frac{\partial \bar{m}_i}{\partial G_i} &= \frac{c_i N_i U_i^2}{2(\sigma_i^2 + 2w_i)^2} \left(\frac{(\sigma_i^2 + 2w_i)\sigma_i^2}{2(w_i + \sigma_i^2)} - \sigma_i^2 - 2w_i \right) \sqrt{\frac{b_{ii}}{2\pi}} \\ &= \frac{c_i N_i U_i^2}{2(\sigma_i^2 + 2w_i)} \left(\frac{\sigma_i^2}{2(\sigma_i^2 + w_i)} - 1 \right) \sqrt{\frac{b_{ii}}{2\pi}} = -\frac{c_i N_i U_i^2 b_{ii}}{2} \sqrt{\frac{b_{ii}}{2\pi}},\end{aligned}\quad (137)$$

1081 where

$$b_{ij} = \frac{1}{w_i + w_j + \sigma_i^2 + \sigma_j^2}. \quad (138)$$

1082 The average fitness for species i is

$$\bar{m}_i = R_i - \frac{a_i}{2} \left((\bar{x}_i - \theta_i)^2 + G_i + \eta_i \right) - c_i \sum_{j=1}^S N_j U_i U_j \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2}. \quad (139)$$

1083 Thus,

$$\frac{\partial \bar{m}_i}{\partial \bar{x}_i} = a_i(\theta_i - \bar{x}_i) - c_i \sum_j N_j U_i U_j b_{ij} (\bar{x}_j - \bar{x}_i) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2}, \quad (140)$$

$$\frac{\partial \bar{m}_i}{\partial G_i} = -\frac{a_i}{2} + \frac{c_i}{2} \sum_{j=1}^S N_j U_i U_j b_{ij} (1 - b_{ij}(\bar{x}_i - \bar{x}_j)^2) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2}. \quad (141)$$

1084 In particular

$$\frac{\partial \bar{m}_i}{\partial G_i} - \frac{\overline{\partial m}_i}{\partial G_i} = -\frac{a_i}{2} + \frac{c_i}{2} \left(N_i U_i^2 b_{ii} \sqrt{\frac{b_{ii}}{2\pi}} + \sum_{j=1}^S N_j U_i U_j b_{ij} (1 - b_{ij}(\bar{x}_i - \bar{x}_j)^2) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2} \right). \quad (142)$$

1085 Applying equations (30a), (40a) and (40b) of the main text recovers system (45) of the main text.

1086 5.8 The relation between competition coefficients and selection

1087 5.8.1 Derivation of analytical approximations

1088 Just as with most calculations in this work, the derivations are straightforward applications of Gaussian
1089 products. That is, if

$$f_1(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(\mu_1 - x)^2}{2\sigma_1^2}\right), \quad f_2(x) = \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(\mu_2 - x)^2}{2\sigma_2^2}\right), \quad (143)$$

1090 then

$$f_1(x)f_2(x) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left(-\frac{(\mu_1 - \mu_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right) \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}} \exp\left(-\frac{(\tilde{\mu} - x)^2}{2\tilde{\sigma}^2}\right), \quad (144)$$

1091 where

$$\tilde{\mu} = \frac{\sigma_2^2 \mu_1 + \sigma_1^2 \mu_2}{\sigma_1^2 + \sigma_2^2}, \quad \tilde{\sigma}^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}. \quad (145)$$

1092 **5.8.1.1 Caclulating $\text{Cov}_{f_{\bar{X}}}(\alpha, \gamma)$**

1093 Recalling

$$\alpha(\bar{x}_i, \bar{x}_j) = \frac{c}{\bar{r}} \sqrt{\frac{b}{2\pi}} \exp\left(-\frac{b}{2} (\bar{x}_i - \bar{x}_j)^2\right), \quad (146)$$

$$\gamma(\bar{x}_i, \bar{x}_j) = cNb \left(1 - b(\bar{x}_i - \bar{x}_j)^2\right) \sqrt{\frac{b}{2\pi}} \exp\left(-\frac{b}{2} (\bar{x}_i - \bar{x}_j)^2\right), \quad (147)$$

1094 we have

$$\begin{aligned} \bar{\alpha} &= \int_{\mathbb{R}} \int_{\mathbb{R}} \alpha(\bar{x}_i, \bar{x}_j) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\ &= \frac{c}{\bar{r}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi(b^{-1} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(b^{-1} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j = \frac{c/\bar{r}}{\sqrt{2\pi(b^{-1} + 2V_{\bar{X}})}}, \end{aligned} \quad (148)$$

$$\begin{aligned} \bar{\gamma} &= \int_{\mathbb{R}} \int_{\mathbb{R}} \gamma(\bar{x}_i, \bar{x}_j) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\ &= cNb \int_{\mathbb{R}} \left\{ 1 - \left[\left(\frac{\bar{x} + bV_{\bar{X}}\bar{x}_j}{1 + bV_{\bar{X}}} - \bar{x}_j \right)^2 + \frac{V_{\bar{X}}}{1 + bV_{\bar{X}}} \right] \right\} \frac{1}{\sqrt{2\pi(b^{-1} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(b^{-1} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j \\ &= cNb \int_{\mathbb{R}} \left\{ 1 - \left[\left(\frac{\bar{x} - \bar{x}_j}{1 + bV_{\bar{X}}} \right)^2 + \frac{V_{\bar{X}}}{1 + bV_{\bar{X}}} \right] \right\} \frac{1}{\sqrt{2\pi(b^{-1} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(b^{-1} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j \\ &= cNb \left(1 - \frac{(1 + bV_{\bar{X}})V_{\bar{X}}}{1 + 2bV_{\bar{X}}} \frac{1}{(1 + bV_{\bar{X}})^2} - \frac{V_{\bar{X}}}{1 + bV_{\bar{X}}} \right) \frac{1}{\sqrt{2\pi(b^{-1} + 2V_{\bar{X}})}} \\ &= cNb \left[1 - \left(\frac{1}{1 + 2bV_{\bar{X}}} + 1 \right) \frac{V_{\bar{X}}}{1 + bV_{\bar{X}}} \right] \frac{1}{\sqrt{2\pi(b^{-1} + 2V_{\bar{X}})}} \\ &= cNb \left(1 - \frac{2V_{\bar{X}}}{1 + 2bV_{\bar{X}}} \right) \sqrt{\frac{b}{2\pi(1 + 2bV_{\bar{X}})}}, \end{aligned} \quad (149)$$

$$\begin{aligned} \text{Var}_{f_{\bar{X}}}(\alpha) &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{\alpha} - \alpha(\bar{x}_i, \bar{x}_j))^2 f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\ &= \frac{c^2}{\bar{r}^2} \left(\sqrt{\frac{b}{4\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} \sqrt{\frac{b}{\pi}} \exp(-b(\bar{x}_i - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j - \frac{1}{2\pi(b^{-1} + 2V_{\bar{X}})} \right) \\ &= \frac{c^2}{\bar{r}^2} \left(\sqrt{\frac{b}{4\pi}} \int_{\mathbb{R}} \sqrt{\frac{1}{2\pi(\frac{1}{2b} + V_{\bar{X}})}} \exp(-b(\bar{x} - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j - \frac{1}{2\pi(b^{-1} + 2V_{\bar{X}})} \right) \\ &= \frac{c^2}{\bar{r}^2} \left(\sqrt{\frac{b}{4\pi}} \sqrt{\frac{1}{2\pi(\frac{1}{2b} + 2V_{\bar{X}})}} - \frac{1}{2\pi(b^{-1} + 2V_{\bar{X}})} \right) = \frac{c^2 b}{2\pi \bar{r}^2} \left(\frac{1}{\sqrt{1 + 4bV_{\bar{X}}}} - \frac{1}{1 + 2bV_{\bar{X}}} \right), \end{aligned} \quad (150)$$

$$\begin{aligned}
\text{Cov}_{f_{\bar{X}}}(\alpha, \gamma) &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{\alpha} - \alpha(\bar{x}_i, \bar{x}_j)) (\bar{\gamma} - \gamma(\bar{x}_i, \bar{x}_j)) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \frac{c^2 N b}{2\bar{r}} \sqrt{\frac{b}{\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} (1 - b(\bar{x}_i - \bar{x}_j)^2) \sqrt{\frac{b}{\pi}} \exp(-b(\bar{x}_i - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j - \bar{\alpha} \bar{\gamma} \\
&= \frac{c^2 N b}{2\bar{r}} \sqrt{\frac{b}{\pi}} \frac{1 - 2bV_{\bar{X}}}{\sqrt{2\pi((2b)^{-1} + 2V_{\bar{X}})}} - \frac{c^2 N b}{\bar{r}} \frac{1 - 2bV_{\bar{X}}}{2\pi(b^{-1} + 2V_{\bar{X}})} \\
&= \frac{c^2 b^2 N}{2\pi\bar{r}} (1 - 2bV_{\bar{X}}) \left(\frac{1}{\sqrt{1 + 4bV_{\bar{X}}}} - \frac{1}{1 + 2bV_{\bar{X}}} \right). \quad (151)
\end{aligned}$$

1095 **5.8.1.2 Caclulating $\text{Cov}_{f_{\bar{X}}}(\alpha, |\beta|)$**

1096 To calculate moments of $|\beta|$ we note that, as a random variable, $|\beta|$ takes a folded normal distribution.
1097 Setting $\Phi(x)$ equal to the cumulative density function of the standard normal distribution and using the
1098 properties of the folded normal distribution, we find

$$|\bar{\beta}| = \sqrt{\frac{2\text{Var}_{f_{\bar{X}}}(\beta)}{\pi}} \exp\left(-\frac{\bar{\beta}^2}{2\text{Var}_{f_{\bar{X}}}(\beta)}\right) - \bar{\beta} \left[1 - 2\Phi\left(\frac{\bar{\beta}}{\sqrt{\text{Var}_{f_{\bar{X}}}(\beta)}}\right)\right] \quad (152)$$

$$\text{Var}_{f_{\bar{X}}}(|\beta|) = \bar{\beta}^2 + \text{Var}_{f_{\bar{X}}}(\beta) - |\bar{\beta}|^2. \quad (153)$$

1099 Recall that

$$\beta(\bar{x}_i, \bar{x}_j) = cNb(\bar{x}_i - \bar{x}_j) \sqrt{\frac{b}{2\pi}} \exp\left(-\frac{b}{2}(\bar{x}_i - \bar{x}_j)^2\right) \quad (154)$$

1100 and hence

$$\begin{aligned}
\bar{\beta} &= \int_{\mathbb{R}} \int_{\mathbb{R}} \beta(\bar{x}_i, \bar{x}_j) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= cNb \int_{\mathbb{R}} (\bar{x} - \bar{x}_j) \frac{1}{\sqrt{2\pi(b^{-1} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(b^{-1} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j = 0, \quad (155)
\end{aligned}$$

$$\begin{aligned}
\text{Var}_{f_{\bar{X}}}(\beta) &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{\beta} - \beta(\bar{x}_i, \bar{x}_j))^2 f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} c^2 N^2 b^2 (\bar{x}_i - \bar{x}_j)^2 \frac{b}{2\pi} \exp(-b(\bar{x}_i - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \sqrt{\frac{b}{4\pi}} c^2 N^2 b^2 \int_{\mathbb{R}} \left[\left(\frac{\bar{x} + 2bV_{\bar{X}}\bar{x}_j}{1 + 2bV_{\bar{X}}} - \bar{x}_j \right)^2 + \frac{V_{\bar{X}}}{1 + 2bV_{\bar{X}}} \right] \frac{\exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(\frac{1}{2b} + V_{\bar{X}})}\right)}{\sqrt{2\pi(\frac{1}{2b} + V_{\bar{X}})}} f_{\bar{X}}(\bar{x}_j) d\bar{x}_j \\
&= \sqrt{\frac{b}{4\pi}} c^2 N^2 b^2 \int_{\mathbb{R}} \left[\frac{(\bar{x} - \bar{x}_j)^2}{(1 + 2bV_{\bar{X}})^2} + \frac{V_{\bar{X}}}{1 + 2bV_{\bar{X}}} \right] \frac{\exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(\frac{1}{2b} + V_{\bar{X}})}\right)}{\sqrt{2\pi(\frac{1}{2b} + V_{\bar{X}})}} f_{\bar{X}}(\bar{x}_j) d\bar{x}_j \\
&= \sqrt{\frac{b}{4\pi}} c^2 N^2 b^2 \left[\frac{(1 + 2bV_{\bar{X}})V_{\bar{X}}}{1 + 4bV_{\bar{X}}} \frac{1}{(1 + 2bV_{\bar{X}})^2} + \frac{V_{\bar{X}}}{1 + 2bV_{\bar{X}}} \right] \frac{1}{\sqrt{2\pi(\frac{1}{2b} + 2V_{\bar{X}})}} \\
&= \frac{b}{\pi} \frac{c^2 N^2 b^2}{\sqrt{1 + 4bV_{\bar{X}}}} \frac{V_{\bar{X}}}{1 + 2bV_{\bar{X}}} \left(\frac{1}{1 + 4bV_{\bar{X}}} + 1 \right) = \frac{2c^2 N^2 b^3 V_{\bar{X}}}{\pi(1 + 4bV_{\bar{X}})^{3/2}}. \quad (156)
\end{aligned}$$

¹¹⁰¹ Thus, using properties of the folded normal distribution, we find

$$\overline{|\beta|} = \sqrt{\frac{2}{\pi}} \frac{cNb^{3/2}}{(1 + 4bV_{\bar{X}})^{3/4}} \sqrt{\frac{2V_{\bar{X}}}{\pi}} = \frac{2}{\pi} \frac{cNb^{3/2}}{(1 + 4bV_{\bar{X}})^{3/4}} \sqrt{V_{\bar{X}}}, \quad (157)$$

$$\text{Var}_{f_{\bar{X}}}(|\beta|) = \frac{c^2 N^2 b^3}{(1 + 4bV_{\bar{X}})^{3/2}} \frac{2V_{\bar{X}}}{\pi} \left(1 - \frac{2}{\pi} \right). \quad (158)$$

¹¹⁰² We also calculate

$$\begin{aligned}
\text{Cov}_{f_{\bar{X}}}(\alpha, \beta) &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{\alpha} - \alpha(\bar{x}_i, \bar{x}_j)) (\bar{\beta} - \beta(\bar{x}_i, \bar{x}_j)) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \frac{c^2 Nb}{2\bar{r}} \sqrt{\frac{b}{\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{x}_i - \bar{x}_j) \sqrt{\frac{b}{\pi}} \exp(-b(\bar{x}_i - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j = 0. \quad (159)
\end{aligned}$$

¹¹⁰³ In attempt to calculate $\text{Cov}_{f_{\bar{X}}}(\alpha, |\beta|)$ we find

$$\begin{aligned}
\text{Cov}_{f_{\bar{X}}}(\alpha, |\beta|) &= \int_{\mathbb{R}} \int_{\mathbb{R}} \alpha(\bar{x}_i, \bar{x}_j) |\beta(\bar{x}_i, \bar{x}_j)| f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j - \bar{\alpha} \overline{|\beta|} \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{c}{\bar{r}} \sqrt{\frac{b}{2\pi}} \exp\left(-\frac{b}{2}(\bar{x}_i - \bar{x}_j)^2\right) cNb |\bar{x}_i - \bar{x}_j| \sqrt{\frac{b}{2\pi}} \exp\left(-\frac{b}{2}(\bar{x}_i - \bar{x}_j)^2\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j - \bar{\alpha} \overline{|\beta|} \\
&= \frac{c^2 Nb}{\bar{r}} \sqrt{\frac{b}{4\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} |\bar{x}_i - \bar{x}_j| \sqrt{\frac{b}{\pi}} \exp(-b(\bar{x}_i - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j - \bar{\alpha} \overline{|\beta|}. \quad (160)
\end{aligned}$$

¹¹⁰⁴ Just as we used the folded normal to find $\overline{|\beta|}$ and $\text{Var}_{f_{\bar{X}}}(|\beta|)$, we can calculate $\text{Cov}_{f_{\bar{X}}}(\alpha, |\beta|)$ by considering

$$\int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{x}_i - \bar{x}_j) \sqrt{\frac{b}{\pi}} \exp(-b(\bar{x}_i - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j = 0 \quad (161)$$

¹¹⁰⁵ and

$$\begin{aligned}
& \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{x}_i - \bar{x}_j)^2 \frac{b}{\pi} \exp(-2b(\bar{x}_i - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \sqrt{\frac{2b}{\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{x}_i - \bar{x}_j)^2 \frac{1}{\sqrt{2\pi^{\frac{1}{4b}}}} \exp\left(-\frac{(\bar{x}_i - \bar{x}_j)^2}{2^{\frac{1}{4b}}}\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \sqrt{\frac{2b}{\pi}} \int_{\mathbb{R}} \left[\left(\frac{\bar{x} + 4bV_{\bar{X}}\bar{x}_j}{1 + 4bV_{\bar{X}}} - \bar{x}_j \right)^2 + \frac{V_{\bar{X}}}{1 + 4bV_{\bar{X}}} \right] \frac{1}{\sqrt{2\pi(\frac{1}{4b} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(\frac{1}{4b} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j \\
&= \sqrt{\frac{2b}{\pi}} \int_{\mathbb{R}} \left[\left(\frac{\bar{x} - \bar{x}_j}{1 + 4bV_{\bar{X}}} \right)^2 + \frac{V_{\bar{X}}}{1 + 4bV_{\bar{X}}} \right] \frac{1}{\sqrt{2\pi(\frac{1}{4b} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(\frac{1}{4b} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j \\
&= \sqrt{\frac{2b}{\pi}} \left[\frac{(1 + 4bV_{\bar{X}})V_{\bar{X}}}{1 + 8bV_{\bar{X}}} \frac{1}{(1 + 4bV_{\bar{X}})^2} + \frac{V_{\bar{X}}}{1 + 4bV_{\bar{X}}} \right] \frac{1}{\sqrt{2\pi(\frac{1}{4b} + 2V_{\bar{X}})}} \\
&= \sqrt{\frac{2b}{\pi}} \frac{2V_{\bar{X}}}{1 + 8bV_{\bar{X}}} \sqrt{\frac{4b}{2\pi(1 + 8bV_{\bar{X}})}} = \frac{b}{\pi} \frac{4V_{\bar{X}}}{(1 + 8bV_{\bar{X}})^{3/2}}. \quad (162)
\end{aligned}$$

¹¹⁰⁶ Hence

$$\int_{\mathbb{R}} \int_{\mathbb{R}} |\bar{x}_i - \bar{x}_j| \sqrt{\frac{b}{\pi}} \exp(-b(\bar{x}_i - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j = \sqrt{\frac{2}{\pi}} \sqrt{\frac{b}{\pi} \frac{4V_{\bar{X}}}{(1 + 8bV_{\bar{X}})^{3/2}}} = \frac{2}{\pi} \frac{\sqrt{2bV_{\bar{X}}}}{(1 + 8bV_{\bar{X}})^{3/2}} \quad (163)$$

¹¹⁰⁷ and

$$\begin{aligned}
\text{Cov}_{f_{\bar{X}}}(\alpha, |\beta|) &= \frac{c^2 N b}{\bar{r}} \sqrt{\frac{b}{4\pi}} \frac{2}{\pi} \frac{\sqrt{2bV_{\bar{X}}}}{(1 + 8bV_{\bar{X}})^{3/4}} - \bar{\alpha} |\beta| \\
&= \frac{2c^2 N b^2}{\pi \bar{r} (1 + 8bV_{\bar{X}})^{3/4}} \sqrt{\frac{V_{\bar{X}}}{2\pi}} - \frac{c}{\bar{r}} \sqrt{\frac{b}{2\pi(1 + 2bV_{\bar{X}})}} \frac{2}{\pi} \frac{c N b^{3/2}}{(1 + 4bV_{\bar{X}})^{3/4}} \sqrt{V_{\bar{X}}} \\
&= \frac{2c^2 N b^2}{\pi \bar{r} (1 + 8bV_{\bar{X}})^{3/4}} \sqrt{\frac{V_{\bar{X}}}{2\pi}} - \frac{2c^2 N b^2}{\pi \bar{r} (1 + 4bV_{\bar{X}})^{3/4}} \sqrt{\frac{V_{\bar{X}}}{2\pi(1 + 2bV_{\bar{X}})}} \\
&= \frac{2c^2 N b^2}{\pi \bar{r}} \sqrt{\frac{V_{\bar{X}}}{2\pi}} \left(\frac{1}{(1 + 8bV_{\bar{X}})^{3/4}} - \frac{1}{(1 + 4bV_{\bar{X}})^{3/4}(1 + 2bV_{\bar{X}})^{1/2}} \right). \quad (164)
\end{aligned}$$

¹¹⁰⁸ **5.8.1.3 Starting the calculation of $\text{Cov}_{f_{\bar{X}}}(\alpha, \mathfrak{C})$**

¹¹⁰⁹ We have

$$\mathfrak{C}(\bar{x}_i, \bar{x}_j) = c^2 N^2 b^2 \left(|\bar{x}_i - \bar{x}_j| + |1 - b(\bar{x}_i - \bar{x}_j)^2| \right)^2 \exp\left(-\frac{b}{2}(\bar{x}_i - \bar{x}_j)^2\right). \quad (165)$$

¹¹¹⁰ Note that the random variable $\delta = \bar{x}_i - \bar{x}_j$ is a mean zero Gaussian random variable with variance $2V_{\bar{X}}$. We
¹¹¹¹ write the probability density function of δ as $f_{\Delta}(\delta)$. Substituting in δ , we can write

$$\begin{aligned}\mathfrak{C}(\delta, 0) &= c^2 N^2 b^2 \left(|\delta| + |1 - b\delta^2| \right)^2 \exp \left(-\frac{b}{2} \delta^2 \right) \\ &= c^2 N^2 b^2 \left(\delta^2 + 2|\delta| - b|\delta|^3 + (1 - b\delta^2)^2 \right) \exp \left(-\frac{b}{2} \delta^2 \right).\end{aligned}\quad (166)$$

From this expression, we see properties of the folded normal distribution can be used to calculate several components of the integral $\text{Cov}_{f_{\bar{X}}}(\alpha, \mathfrak{C})$, but a major technical challenge lies in calculating

$$\int_{\mathbb{R}} | |\delta| - b|\delta|^3 | \exp \left(-\frac{b}{2} \delta^2 \right) f_{\Delta}(\delta) d\delta.\quad (167)$$

Instead of overcoming this challenge to find an analytical form of $\text{Cov}_{f_{\bar{X}}}(\alpha, \mathfrak{C})$ we turn to a numerical approach outlined in the following section.

5.8.2 Numerical estimates for heterogeneous N and G

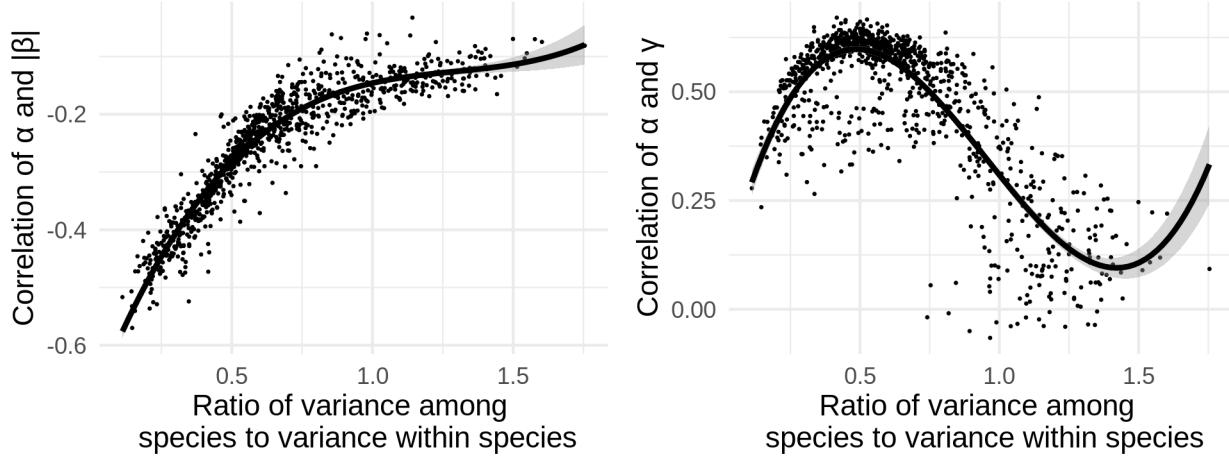


Figure 7: Numerical estimate for the correlations of selection gradients and competition coefficients.

Details on simulations, table of parameters, distributions of a and c .

1118 **References**

- 1119 Abdala-Roberts, Luis, and Kailen A. Mooney. 2014. “Ecological and Evolutionary Consequences of Plant
1120 Genotype Diversity in a Tri-Trophic System.” *Ecology* 95 (10). Wiley: 2879–93.
- 1121 Ågren, Göran I., and Folke O. Andersson. 2012. *Terrestrial Ecosystem Ecology: Principles and Applications*.
1122 Cambridge University Press.
- 1123 Barton, N.H., A.M. Etheridge, and A. Véber. 2017. “The Infinitesimal Model: Definition, Derivation, and
1124 Implications.” *Theoretical Population Biology* 118 (December). Elsevier BV: 50–73.
- 1125 Barton, Nick, and Alison Etheridge. 2019. “Mathematical Models in Population Genetics.” Wiley.
- 1126 Barton, Nick, Alison Etheridge, and Amandine Véber. 2013. “Modelling Evolution in a Spatial Continuum.”
1127 *Journal of Statistical Mechanics: Theory and Experiment* 2013 (01). IOP Publishing: P01002.
- 1128 Bertoin, Jean, and Jean-François Le Gall. 2003. “Stochastic Flows Associated to Coalescent Processes.”
1129 *Probability Theory and Related Fields* 126 (2). Springer Science; Business Media: 261–88.
- 1130 Bulmer. 1980. *The Mathematical Theory of Quantitative Genetics*. Oxford University Press.
- 1131 Bürger, Reinhard. 1986. “On the Maintenance of Genetic Variation: Global Analysis of Kimuras Continuum-
1132 of-Alleles Model.” *Journal of Mathematical Biology* 24 (3). Springer Nature: 341–51.
- 1133 ———. 2000. *The Mathematical Theory of Selection, Recombination, and Mutation*. Wiley.
- 1134 Champagnat, Nicolas, Régis Ferrière, and Sylvie Méléard. 2006. “Unifying Evolutionary Dynamics: From
1135 Individual Stochastic Processes to Macroscopic Models.” *Theoretical Population Biology* 69 (3). Elsevier
1136 BV: 297–321.
- 1137 Chesson, Peter. 2000. “Mechanisms of Maintenance of Species Diversity.” *Annual Review of Ecology and
1138 Systematics* 31 (1). Annual Reviews: 343–66.
- 1139 Conner, Jeffrey K. 2004. *A Primer of Ecological Genetics*. Sinauer Associates.
- 1140 Crow, James F., and Motoo Kimura. 1970. *An Introduction to Population Genetics Theory*. The Blackburn
1141 Press.
- 1142 Crutsinger, Gregory M. 2015. “A Community Genetics Perspective: Opportunities for the Coming Decade.”
1143 *New Phytologist* 210 (1). Wiley: 65–70.
- 1144 Da Prato, Giuseppe, and Jerzy Zabczyk. 2014. *Stochastic Equations in Infinite Dimensions*. Cambridge
1145 University Press.
- 1146 David, Thomas I., Jonathan Storkey, and Carly J. Stevens. 2019. “Understanding How Changing Soil Ni-
1147 tragen Affects Plantpollinator Interactions.” *Arthropod-Plant Interactions* 13 (5). Springer Science; Business
1148 Media LLC: 671–84.
- 1149 Dawson, Donald A. 1975. “Stochastic Evolution Equations and Related Measure Processes.” *Journal of
1150 Multivariate Analysis* 5 (1). Elsevier BV: 1–52.
- 1151 ———. 1993. “Measure-Valued Markov Processes.” In *École d’été de Probabilités de Saint-Flour Xxi-1991*,
1152 1–260. Springer.
- 1153 Débarre, Florence, Sam Yeaman, and Frédéric Guillaume. 2015. “Evolution of Quantitative Traits Under a
1154 Migration-Selection Balance: When Does Skew Matter?” *The American Naturalist* 186 Suppl 1: S37–47.
- 1155 Doebeli, Michael. 1996. “Quantitative Genetics and Population Dynamics.” *Evolution* 50 (2). Wiley:
1156 532–46.
- 1157 Etheridge, Alison M. 2000. *An Introduction to Superprocesses*. American Mathematical Society.
- 1158 ———. 2008. “Drift, Draft and Structure: Some Mathematical Models of Evolution.” In *Stochastic Models
1159 in Biological Sciences*. Institute of Mathematics Polish Academy of Sciences.

- 1160 Etheridge, Alison, and Peter March. 1991. "A Note on Superprocesses." *Probability Theory and Related*
1161 *Fields* 89 (2). Springer: 141–47.
- 1162 Evans, Lawrence C. 2010. *Partial Differential Equations: Second Edition*. American Mathematical Society.
- 1163 ———. 2014. *An Introduction to Stochastic Differential Equations*. American Mathematical Society.
- 1164 Ewens, Warren J. 2004. *Mathematical Population Genetics*. Springer New York.
- 1165 Falconer, Kenneth. 2014. *Fractal Geometry: Mathematical Foundations and Applications*. Wiley.
- 1166 Farlow, Stanley J. 1993. *Partial Differential Equations for Scientists and Engineers*. Dover.
- 1167 Feller, William. 1951. "Diffusion Processes in Genetics." In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 227–46. University of California Press.
- 1168 Felsenstein, Joseph. 1975. "A Pain in the Torus: Some Difficulties with Models of Isolation by Distance." *The American Naturalist* 109 (967). University of Chicago Press: 359–68.
- 1169 Fisher, R. A. 1923. "XXI.-on the Dominance Ratio." *Proceedings of the Royal Society of Edinburgh* 42.
- 1170 Cambridge University Press: 321–41.
- 1171 Fitzpatrick, Connor R., Anurag A. Agrawal, Nathan Basiliko, Amy P. Hastings, Marney E. Isaac, Michael
1172 Preston, and Marc T. J. Johnson. 2015. "The Importance of Plant Genotype and Contemporary Evolution
1173 for Terrestrial Ecosystem Processes." *Ecology* 96 (10). Wiley: 2632–42.
- 1174 Fitzpatrick, Connor R., Anna V. Mikhailichenko, Daniel N. Anstett, and Marc T. J. Johnson. 2017. "The
1175 Influence of Range-Wide Plant Genetic Variation on Soil Invertebrate Communities." *Ecography* 41 (7).
- 1176 Wiley: 1135–46.
- 1177 Frank, S. A. 2012. "Natural Selection. IV. The Price Equation." *Journal of Evolutionary Biology* 25 (6).
- 1178 Wiley: 1002–19.
- 1179 Fridley, Jason D. 2017. "Plant Energetics and the Synthesis of Population and Ecosystem Ecology." Edited
1180 by David Gibson. *Journal of Ecology* 105 (1). Wiley: 95–110.
- 1181 Fussman, G. F., M. Loreau, and P. A. Abrams. 2007. "Eco-Evolutionary Dynamics of Communities and
1182 Ecosystems." *Functional Ecology* 21 (3). Wiley: 465–77.
- 1183 Guimarães, Paulo R., Pedro Jordano, and John N. Thompson. 2011. "Evolution and Coevolution in Mutualistic
1184 Networks." *Ecology Letters* 14 (9). Wiley: 877–85.
- 1185 Guimarães, Paulo R., Mathias M. Pires, Pedro Jordano, Jordi Bascompte, and John N. Thompson. 2017.
1186 "Indirect Effects Drive Coevolution in Mutualistic Networks." *Nature* 550 (7677). Springer Science; Business
1187 Media LLC: 511–14.
- 1188 Harmon, Luke J., Cecilia S. Andreazza, Florence Débarre, Jonathan Drury, Emma E. Goldberg, Ayana B.
1189 Martins, Carlos J. Melián, et al. 2019. "Detecting the Macroevolutionary Signal of Species Interactions."
1190 *Journal of Evolutionary Biology* 32 (8). Wiley: 769–82.
- 1191 Harte, John. 2011. *Maximum Entropy and Ecology*. Oxford University Press.
- 1192 Harte, John, and Erica A. Newman. 2014. "Maximum Information Entropy: A Foundation for Ecological
1193 Theory." *Trends in Ecology & Evolution* 29 (7). Elsevier BV: 384–89.
- 1194 Hickerson, M.J., B.C. Carstens, J. Cavender-Bares, K.A. Crandall, C.H. Graham, J.B. Johnson, L. Rissler,
1195 P.F. Victoriano, and A.D. Yoder. 2010. "Phylogeography's Past, Present, and Future: 10 Years After Avise,
1196 2000." *Molecular Phylogenetics and Evolution* 54 (1). Elsevier BV: 291–301.
- 1197 Hofbauer, Josef, and Karl Sigmund. 1998. *Evolutionary Games and Population Dynamics*. Cambridge
1198 University Press.
- 1199 Holt, Robert D. 1987. "On the Relation Between Niche Overlap and Competition: The Effect of Incommen-
1200 surable Niche Dimensions." *Oikos* 48 (1). JSTOR: 110.

- 1203 Kendall, David G. 1966. "Branching Processes Since 1873." *Journal of the London Mathematical Society*
 1204 s1-41 (1). Wiley: 385–406.
- 1205 Kimmel, Marek, and David E. Axelrod. 2015. *Branching Processes in Biology*. Springer New York.
- 1206 Kimura, M. 1965. "A Stochastic Model Concerning the Maintenance of Genetic Variability in Quantitative
 1207 Characters." *Proceedings of the National Academy of Sciences* 54 (3). Proceedings of the National Academy
 1208 of Sciences: 731–36.
- 1209 Kimura, M., and J. F. Crow. 1978. "Effect of Overall Phenotypic Selection on Genetic Change at Individual
 1210 Loci." *Proceedings of the National Academy of Sciences* 75 (12). Proceedings of the National Academy of
 1211 Sciences: 6168–71.
- 1212 Kirkpatrick, Mark, Toby Johnson, and Nick Barton. 2002. "General Models of Multilocus Evolution." *Genetics* 161 (4). Genetics: 1727–50.
- 1214 Kolmogorov, A.N., and S.V. Fomin. 1999. *Elements of the Theory of Functions and Functional Analysis*. v.
 1215 1. Dover.
- 1216 Konno, N., and T. Shiga. 1988. "Stochastic Partial Differential Equations for Some Measure-Valued Diffusions." *Probability Theory and Related Fields* 79 (2). Springer Nature: 201–25.
- 1218 Kopp, Michael, and Sergey Gavrilets. 2006. "Multilocus Genetics and the Coevolution of Quantitative Traits." *Evolution* 60 (7). Wiley: 1321–36.
- 1220 Kölzsch, Andrea, Adriana Alzate, Frederic Bartumeus, Monique de Jager, Ellen J. Weerman, Geerten M.
 1221 Hengeveld, Marc Naguib, Bart A. Nolet, and Johan van de Koppel. 2015. "Experimental Evidence for
 1222 Inherent Lévy Search Behaviour in Foraging Animals." *Proceedings of the Royal Society B: Biological Sciences*
 1223 282 (1807). The Royal Society: 20150424.
- 1224 Kraft, Nathan J. B., William K. Cornwell, Campbell O. Webb, and David D. Ackerly. 2007. "Trait Evolution,
 1225 Community Assembly, and the Phylogenetic Structure of Ecological Communities." *The American Naturalist*
 1226 170 (2). University of Chicago Press: 271–83.
- 1227 Krylov, N. V., and B. L. Rozovskii. 1981. "Stochastic Evolution Equations." *Journal of Soviet Mathematics*
 1228 16 (4). Springer Science; Business Media LLC: 1233–77.
- 1229 Lande, Russel. 1976. "Natural Selection and Random Genetic Drift in Phenotypic Evolution." *Evolution* 30
 1230 (2). Wiley: 314–34.
- 1231 ———. 1980. "The Genetic Covariance between Characters Maintained by Pleiotropic Mutations." *Genetics*
 1232 94 (1): 203–15.
- 1233 Lande, Russell. 1982. "A Quantitative Genetic Theory of Life History Evolution." *Ecology* 63 (3). Wiley:
 1234 607–15.
- 1235 Lande, Russell, and Stevan J. Arnold. 1983. "The Measurement of Selection on Correlated Characters." *Evolution* 37 (6). JSTOR: 1210.
- 1237 Lande, Russell, Steinar Engen, and Bernt-Erik Sæther. 2003. "Demographic and Environmental Stochasticity." In *Stochastic Population Dynamics in Ecology and Conservation*, 1–24. Oxford University Press.
- 1239 Landis, Michael J., and Joshua G. Schraiber. 2017. "Pulsed Evolution Shaped Modern Vertebrate Body
 1240 Sizes." *Proceedings of the National Academy of Sciences* 114 (50). Proceedings of the National Academy of
 1241 Sciences: 13224–9.
- 1242 Levins, Richard. 1968. *Evolution in Changing Environments: Some Theoretical Explorations.* (MPB-2)
 1243 (Monographs in Population Biology). Princeton University Press.
- 1244 Li, Zeng-Hu. 1998. "Absolute Continuity of Measure Branching Processes with Interaction." *Chinese Journal
 1245 of Applied Probability and Statistics* 14. Citeseer: 231–42.
- 1246 Lion, Sébastien. 2018. "Theoretical Approaches in Evolutionary Ecology: Environmental Feedback as a
 1247 Unifying Perspective." *The American Naturalist* 191 (1). University of Chicago Press: 21–44.

- 1248 Loreau, Michel. 2010. *From Populations to Ecosystems: Theoretical Foundations for a New Ecological*
1249 *Synthesis*. Princeton University Press.
- 1250 Lynch, Michael, and Bruce Walsh. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates
1251 is an imprint of Oxford University Press.
- 1252 MacArthur, Robert H. 1969. “Species Packing, and what Competition Minimizes.” *Proceedings of the*
1253 *National Academy of Sciences* 64 (4). Proceedings of the National Academy of Sciences: 1369–71.
- 1254 ———. 1970. “Species Packing and Competitive Equilibrium for Many Species.” *Theoretical Population*
1255 *Biology* 1 (1). Elsevier BV: 1–11.
- 1256 ———. 1972. *Geographical Ecology*. Princeton University Press.
- 1257 MacArthur, Robert H., and Richard Levins. 1967. “The Limiting Similarity, Convergence, and Divergence
1258 of Coexisting Species.” *The American Naturalist* 101 (921). University of Chicago Press: 377–85.
- 1259 Manceau, Marc, Amaury Lambert, and Hélène Morlon. 2016. “A Unifying Comparative Phylogenetic
1260 Framework Including Traits Coevolving Across Interacting Lineages.” *Systematic Biology*, December. Oxford
1261 University Press (OUP), syw115.
- 1262 Marx, Hannah E., Cédric Dentant, Julien Renaud, Romain Delunel, David C. Tank, and Sébastien Lavergne.
1263 2017. “Riders in the Sky (Islands): Using a Mega-Phylogenetic Approach to Understand Plant Species
1264 Distribution and Coexistence at the Altitudinal Limits of Angiosperm Plant Life.” *Journal of Biogeography*
1265 44 (11). Wiley: 2618–30.
- 1266 McPeek, Mark A. 2017. *Evolutionary Community Ecology*. Princeton University Press.
- 1267 Meester, Luc De, Kristien I. Brans, Lynn Govaert, Caroline Souffreau, Shinjini Mukherjee, Hélène Vanvelk,
1268 Konrad Korzeniowski, et al. 2018. “Analyzing Eco-Evolutionary Dynamics - the Challenging Complexity of
1269 the Real World.” *Functional Ecology*. Wiley.
- 1270 Méléard, M, and S Roelly. 1992. “Interacting Branching Measure Processes.” *Stochastic Partial Differential*
1271 *Equations and Applications* (G. Da Prato and L. Tubaro, Eds.), 246–56.
- 1272 ———. 1993. “Interacting Measure Branching Processes. Some Bounds for the Support.” *Stochastics and*
1273 *Stochastic Reports* 44 (1-2). Informa UK Limited: 103–21.
- 1274 Mubayi, Anuj, Christopher Kribs, Viswanathan Arunachalam, and Carlos Castillo-Chavez. 2019. “Studying
1275 Complexity and Risk Through Stochastic Population Dynamics: Persistence, Resonance, and Extinction in
1276 Ecosystems.” In *Handbook of Statistics*, 157–93. Elsevier.
- 1277 Nowak, Martin A. 2006. *Evolutionary Dynamics: Exploring the Equations of Life*. Belknap Press.
- 1278 Nuismer, Scott L., Michael Doebeli, and Danny Browning. 2005. “The Coevolutionary Dynamics of An-
1279 tagonistic Interactions Mediated by Quantitative Traits with Evolving Variances.” *Evolution* 59 (10). The
1280 Society for the Study of Evolution: 2073.
- 1281 Nuismer, Scott L., and Luke J. Harmon. 2014. “Predicting Rates of Interspecific Interaction from Phyloge-
1282 netic Trees.” Edited by Jerome Chave. *Ecology Letters* 18 (1). Wiley: 17–27.
- 1283 Nuismer, Scott L., Pedro Jordano, and Jordi Bascompte. 2012. “Coevolution and the Architecture of
1284 Mutualistic Networks.” *Evolution* 67 (2). Wiley: 338–54.
- 1285 Nuismer, Scott L., Benjamin J. Ridenhour, and Benjamin P. Oswald. 2007. “Antagonistic Coevolution
1286 Mediated by Phenotypic Differences Between Quantitative Traits.” *Evolution* 61 (8). Wiley: 1823–34.
- 1287 Nuismer, Scott L., Bob Week, and Marcelo A. Aizen. 2018. “Coevolution Slows the Disassembly of Mu-
1288 tualistic Networks.” *The American Naturalist* 192 (4). University of Chicago Press: 490–502.
- 1289 Nuland, Michael E. Van, Ian M. Ware, Joseph K. Bailey, and Jennifer A. Schweitzer. 2019. “Ecosystem
1290 Feedbacks Contribute to Geographic Variation in Plant-Soil Eco-Evolutionary Dynamics Across a Fertility
1291 Gradient.” Edited by Franziska Brunner. *Functional Ecology* 33 (1). Wiley: 95–106.

- 1292 Page, Karen M., and Martin A. Nowak. 2002. "Unifying Evolutionary Dynamics." *Journal of Theoretical*
1293 *Biology* 219 (1). Elsevier BV: 93–98.
- 1294 Parachnowitsch, Amy L, Jessamyn S Manson, and Nina Sletvold. 2018. "Evolutionary Ecology of Nectar."
1295 *Annals of Botany* 123 (2). Oxford University Press (OUP): 247–61.
- 1296 Patel, Swati, and Reinhard Bürger. 2019. "Eco-Evolutionary Feedbacks Between Prey Densities and Linkage
1297 Disequilibrium in the Predator Maintain Diversity." *Evolution* 73 (8). Wiley: 1533–48.
- 1298 Perkins, Edwin A. 1991. "Conditional Dawson-Watanabe Processes and Fleming-Viot Processes." In *Semi-*
1299 *ninar on Stochastic Processes, 1991*, 143–56. Birkhäuser Boston.
- 1300 ———. 1992. "Measure-Valued Branching Diffusions with Spatial Interactions." *Probability Theory and*
1301 *Related Fields* 94 (2). Springer Science; Business Media LLC: 189–245.
- 1302 ———. 1995. *On the Martingale Problem for Interactive Measure-Valued Branching Diffusions*. Amer
1303 Mathematical Society.
- 1304 Price, George R. 1970. "Selection and Covariance." *Nature* 227 (5257). Springer Nature: 520–21.
- 1305 Queller, David C. 2017. "Fundamental Theorems of Evolution." *The American Naturalist* 189 (4). University
1306 of Chicago Press: 345–53.
- 1307 Reimers, Mark. 1989. "One Dimensional Stochastic Partial Differential Equations and the Branching Mea-
1308 sure Diffusion." *Probability Theory and Related Fields* 81 (3). Springer Nature: 319–40.
- 1309 Robertson, Alan. 1966. "A Mathematical Model of the Culling Process in Dairy Cattle." *Animal Science* 8
1310 (1). Cambridge University Press: 95–108.
- 1311 Roughgarden, Joan. 1979. *Theory of Population Genetics and Evolutionary Ecology: An Introduction*.
1312 Macmillan.
- 1313 Rudman, Seth M., Matthew A. Barbour, Katalin Csilléry, Phillip Gienapp, Frederic Guillaume, Nelson G.
1314 Hairston Jr, Andrew P. Hendry, et al. 2017. "What Genomic Data Can Reveal About Eco-Evolutionary
1315 Dynamics." *Nature Ecology & Evolution* 2 (1). Springer Nature: 9–15.
- 1316 Schreiber, Sebastian J., Swati Patel, and Casey terHorst. 2018. "Evolution as a Coexistence Mechanism:
1317 Does Genetic Architecture Matter?" *The American Naturalist* 191 (3). University of Chicago Press: 407–20.
- 1318 Schuster, Peter, and Karl Sigmund. 1983. "Replicator Dynamics." *Journal of Theoretical Biology* 100 (3).
1319 Elsevier BV: 533–38.
- 1320 Skovmand, Lotte H., Charles C.Y. Xu, Maria R. Servedio, Patrik Nosil, Rowan D.H. Barrett, and Andrew
1321 P. Hendry. 2018. "Keystone Genes." *Trends in Ecology & Evolution* 33 (9). Elsevier BV: 689–700.
- 1322 Sterner, R.W., and J.J. Elser. 2008. "Ecological Stoichiometry: Overview." In *Encyclopedia of Ecology*,
1323 1101–16. Elsevier.
- 1324 Taylor, Peter D., and Leo B. Jonker. 1978. "Evolutionary Stable Strategies and Game Dynamics." *Mathe-*
1325 *matical Biosciences* 40 (1-2). Elsevier BV: 145–56.
- 1326 Tilman, David. 1982. *Resource Competition and Community Structure*. Princeton University Press.
- 1327 Turelli, Michael. 1984. "Heritable Genetic Variation via Mutation-Selection Balance: Lerchs Zeta Meets the
1328 Abdominal Bristle." *Theoretical Population Biology* 25 (2). Elsevier: 138–93.
- 1329 ———. 1986. "Gaussian Versus Non-Gaussian Genetic Analyses of Polygenic Mutation-Selection Balance."
1330 In *Evolutionary Processes and Theory*, 607–28. Academic Press.
- 1331 ———. 2017. "Commentary: Fisher's Infinitesimal Model: A Story for the Ages." *Theoretical Population*
1332 *Biology* 118 (December). Elsevier BV: 46–49.
- 1333 Turelli, Michael, and Nick Barton. 1994. "Genetic and statistical analyses of strong selection on polygenic
1334 traits: what, me normal?" *Genetics* 138 (3): 913–41.

- ¹³³⁵ Walsh, John B. 1986. “An Introduction to Stochastic Partial Differential Equations.” In *Lecture Notes in Mathematics*, 265–439. Springer Berlin Heidelberg.
- ¹³³⁶
- ¹³³⁷ Watanabe, Shinzo. 1968. “A Limit Theorem of Branching Processes and Continuous State Branching Processes.” *Journal of Mathematics of Kyoto University* 8 (1). Duke University Press: 141–67.
- ¹³³⁸
- ¹³³⁹ Wright, Sewall. 1931. “Evolution in Mendelian Populations.” *Genetics* 16 (2). Genetics: 97–159.
- ¹³⁴⁰ Xiao, Xiao, Daniel J. McGlinn, and Ethan P. White. 2015. “A Strong Test of the Maximum Entropy Theory of Ecology.” *The American Naturalist* 185 (3). University of Chicago Press: E70–E80.
- ¹³⁴¹