

A White Noise Approach to Evolutionary Ecology

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

3/1/2020

Abstract

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

1 Introduction

Our goal in this manuscript is to develop a rigorous, but accessible approach to synthesize the stochastic dynamics of abundance, mean trait and heritable variation in biological populations for the study of theoretical evolutionary ecology. A primary aim of theoretical evolutionary ecology is the development of mathematical approaches to describe the evolution of populations and their interactions with both the biotic and abiotic environments in which they are embedded. Given this consideration, a natural scope for such an approach centers on quantifying the abundance dynamics of populations and the evolution of traits mediating their interactions as functions of relevant abiotic factors. Although taking into account abundance, phenotype and environment provides the basis for a partial understanding of the complex nature of biological communities, a deeper understanding must account for the effects of contemporary dispersal and the phylogeographic history of interacting lineages (Kraft et al. 2007; Hickerson et al. 2010; Manceau, Lambert, and Morlon 2016; McPeek 2017) along with the genetic basis of ecologically relevant traits (Conner 2004; Fussman, Loreau, and Abrams 2007) and feedbacks between populations and the biogeochemical cycles they ultimately depend on (Loreau 2010; Ågren and Andersson 2012). It is therefore ideal that the development of any such mathematical approach anticipates extensions to account for these important factors shaping ecological communities, especially as empirical and conceptual work in these directions continues to grow (Abdala-Roberts and Mooney 2014; Kölzsch et al. 2015; Crutsinger 2015; Fitzpatrick et al. 2015, 2017; Marx et al. 2017; Rudman et al. 2017; Skovmand et al. 2018; Nuland et al. 2019; Harmon et al. 2019). Furthermore, the approach would benefit from a stochastic component to capture the chance nature of biological reality (Lande, Engen, and SÆther 2003; Meester et al. 2018; Mubayi et al. 2019) and serve as a basis for the construction of statistical methods that measure evolutionary and ecological processes occurring in the wild. Such methods will tether theory to reality and allow for rigorous tests of hypotheses on the structure and behavior of ecological communities. In this paper we introduce a framework that establishes a formal connection between the continuous-time dynamics of abundance and quantitative traits in stochastically evolving populations, demonstrate its utility through the derivation and analysis of a model of diffuse coevolution and discuss how it can be extended to account for the details mentioned above.

Current theoretical approaches to synthesize evolution and ecology have capitalized on the fact that biological fitness plays a key role in determining both sets of dynamics. While correlation of fitness and genotype is the

49 basis of evolution by natural selection, the mean fitness across all individuals in a population determines
50 the growth, stasis or decline of abundance. In section 2.1 we review the mathematical formalization of
51 this connection, which has been established in the contexts of population genetics (Crow and Kimura 1970;
52 Roughgarden 1979), evolutionary game theory (Hofbauer and Sigmund 1998; Nowak 2006; Lion 2018),
53 quantitative genetics (Lande 1982; Doebeli 1996; Lion 2018) and a unifying framework for these three distinct
54 approaches to evolutionary theory (Champagnat, Ferrière, and Méléard 2006) which is intimately related to
55 the approach we take here.

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

79 To provide motivation for the stochastic equations developed later and background for our model of coevolution,
80 we begin with §2.1 by briefly summarizing derivations of deterministic dynamics of biological populations.
81 Starting with a partial differential equation (PDE), we arrive at a general set of ordinary differential equations
82 modelling the dynamics of abundance, trait mean and trait variance. From this we observe that replacing
83 the PDE with a SPDE provides a path to derive SDE describing the evolutionary response to demographic
84 stochasticity. We accomplish this in §2.2 by introducing a set of mathematical tools based on the calculus of
85 white noise and discuss how a diffusion limit of a spatially structured branching process leads to the natural
86 SPDE appropriate for our study. The diffusion limit in turn provides a rigorous method for constructing
87 fitness functions used in models of evolutionary ecology. We employ these tools to derive a system of
88 SDE generalizing our deterministic results to account for demographic stochasticity. However, although
89 biologically insightful, these equations remain difficult to analyze and implement numerically. In §2.3 we use
90 an assumption of normally distributed trait values to simplify these expressions into formula that are much
91 easier to work with. We then account for the constraint of adaptive evolution on the amount of heritable
92 variation in a population by extending these results via a model of imperfect inheritance. The resulting
93 equations coincide with classical results in quantitative genetics as a special case. We then demonstrate
94 the utility of our approach by combining the derived equations of population dynamics with classical niche
95 theory in §3 to formulate of a model of coevolution across a guild of S species participating in exploitative
96 competition along a common resource continuum. In SM §5.7 we apply a classical theorem on rescaled limits
97 of BBM that allow for ecological interactions to provide a rigorous derivation. To gain biological insight, in
98 §3.2 we numerically integrate our model of coevolution for $S = 100$ species, tracking the dynamics of traits and
99 abundances, under scenarios of weak and strong competition. We include an account of the natural history
100 of the simulated system and discuss the significance of demographic stochasticity for structuring ecological
communities. In §3.3 we provide formula for selection gradients and competition coefficients implied by our

102 model and use these to investigate the relationship between the strengths of competition and coevolution.
103 Finally, §4 concludes with a summary of accomplishments, a few remarks on the limits of this approach
104 and future directions to incorporate more explicitly the genetic architecture of populations, feedbacks with
105 ecosystem processes and the macroevolutionary history of interacting lineages.

106 2 The framework

107 At the core of our approach is a stochastic analog of the replicator equation with mutation in continuous
108 time and phenotypic space (Taylor and Jonker 1978; Schuster and Sigmund 1983). From this stochastic
109 replicator-mutator equation we derive a system of SDE for the dynamics of abundance, mean trait and
110 additive genetic variance of a population. Hence, our approach develops a quantitative genetic theory of
111 evolutionary ecology. A popular alternative to quantitative genetics is the theory of adaptive dynamics. As
112 demonstrated by Page and Nowak (2002), the canonical equation of adaptive dynamics can be derived from
113 the replicator-mutator equation. Thus, one could start from the atomic roots of our approach and pursue a
114 stochastic adaptive dynamic theory instead. We choose the former in anticipation of an extension of our
115 approach that explicitly models the genetic details of populations.

116 In this section we review the derivations of the replicator-mutator equation and trait dynamics from abundance
117 dynamics and extend these formulae along with related results to the stochastic case. The results established
118 in this section provide the framework from which larger scale ecological structures, such as species abundance
119 distributions and interaction networks, can be computed.

120 2.1 Deterministic dynamics

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

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

128 Starting from this model, we get the total abundance of the population as $N = \sum_{i=1}^K \nu_i$, the frequency of
129 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
130 some elementary calculus to derive the dynamics of abundance dN/dt and the dynamics of allele frequencies
131 $dp_1/dt, \dots, dp_K/dt$ as

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

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

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

¹⁴¹ Inspired by equations (1)-(3), we derive an analog of the replicator-mutator equation for a continuum of
¹⁴² types (that is, for a quantitative trait). In particular, we model a continuously reproducing population
¹⁴³ with trait values $x \in \mathbb{R}$ and an abundance density $\nu(x, t)$ that represents the amount of individuals in the
¹⁴⁴ population with trait value x at time t . Hence, the abundance density satisfies $N(t) = \int_{-\infty}^{+\infty} \nu(x, t) dx$ and
¹⁴⁵ $p(x, t) = \nu(x, t)/N(t)$ is the relative density of trait x which we also refer to as the phenotypic distribution.

¹⁴⁶ To stay within the realm of biological plausibility, we require a set technical assumptions. First, we assume
¹⁴⁷ the initial abundance density is continuous, non-negative, integrable and has finite trait mean and variance.
¹⁴⁸ 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 the
¹⁶⁹ 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

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

¹⁷² By our assumptions on mutation and rate of growth, $\nu(x, t)$ is twice differentiable with respect to x and
¹⁷³ $\int_{-\infty}^{+\infty} \nu(x, t) dx < \infty$ for all $t \geq 0$ (SM §??). This implies that we are justified in swapping the order of

¹⁷⁴ 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
¹⁷⁵ theorem of calculus. Biological reasoning agrees with this latter result since mutation neither creates nor
¹⁷⁶ destroys individuals, but merely changes their type from their parental type. Taking the same approach, we
¹⁷⁷ derive the dynamics of the phenotypic distribution $p(x, t)$ in response to selection and mutation as

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

¹⁷⁸ This result closely resembles Kimura's continuum-of-alleles model (Kimura 1965; Bürger 1986). The primary
¹⁷⁹ difference being that our model utilizes diffusion instead of convolution with an arbitrary mutation kernel. Of
¹⁸⁰ course, our model of mutation can be derived as an approximation to Kimura's model, which has been referred
¹⁸¹ to as the Gaussian allelic approximation (in reference to the distribution of mutational effects at a given
¹⁸² locus of a genome on the values of traits, Bürger 2000), the infinitesimal model (in reference to modelling
¹⁸³ continuous traits as being encoded by an infinite number of loci each having infinitesimal effect, Barton,
¹⁸⁴ Etheridge, and Véber 2017) and the Gaussian descendants approximation (in reference to offspring trait
¹⁸⁵ values being normally distributed around their parental values, Turelli 2017). Alternatively, since diffusion is
¹⁸⁶ the continuous-time equivalent to convolution against a Gaussian kernel (SM §5.2), equation (9) can also be
¹⁸⁷ seen as a special case of the continuum-of-alleles model.

¹⁸⁸ The covariance of fitness and phenotype across the population is defined as

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

¹⁸⁹ Hence, the dynamics of the mean trait $\bar{x}(t)$ can be derived as

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

¹⁹⁰ Equation (11) is a continuous time analog of the well known Robertson-Price equation (Robertson 1966; Price
¹⁹¹ 1970; Frank 2012; Queller 2017; Lion 2018). The covariance of fitness and phenotype creates change in \bar{x} to
¹⁹² maximize mean fitness \bar{m} . Since this change is driven by a covariance with respect to phenotypic diversity,
¹⁹³ the response in mean trait to selection is mediated by the phenotypic variance. In particular, when $\sigma^2 = 0$,
¹⁹⁴ \bar{x} will not respond to selection. The second term, due to mutation, represents the so-called *transmission*
¹⁹⁵ *bias* (Frank 2012). When $p(x, t)$ is symmetric around \bar{x} , this term disappears due to equal bias in opposite
¹⁹⁶ directions cancelling the effect of mutation on \bar{x} . Following the approach taken to calculate the evolution of
¹⁹⁷ \bar{x} , we find the response of phenotypic variation to this model of selection and mutation is

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

¹⁹⁸ For the sake of space we relegate the derivation of $d\sigma^2/dt$ to SM §5.3. In the absence of mutation equation
¹⁹⁹ (12) agrees with the result derived by Lion (2018) for discrete phenotypes. From a statistical perspective, if

200 we think of $(x - \bar{x})^2$ as a square error, then in analogy to the dynamics of the mean trait, we see that the
201 response in σ^2 to selection can be expressed as a covariance of fitness and square error, which is defined in
202 analogy to $\text{Cov}_t(m(\nu, x), x)$. This covariance also creates change in σ^2 that maximizes mean fitness \bar{m} . In SM
203 §5.3 we show when $p(x, t)$ is a Gaussian bell curve, the increase in variance due to mutation simplifies to μ .

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

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

228 2.2 White noise calculus and superprocesses

229 2.2.1 Definition and basic properties of white noise

230 One can think of white noise as the static seen on old television sets or infinitely detailed random dust spread
231 across both time and space. From a more mathematical, yet still informal perspective, white noise can be
232 thought of as a stochastic process. That is, we can picture white noise as a collection of random variables
233 indexed by time and possibly space. In relation to Brownian motion, denoted by W , white noise can be
234 interpreted of as the derivative of Brownian motion with respect to time, denoted \dot{W} . Since Brownian motion
235 takes infinitesimally small Gaussian distributed jumps at each time point, this leads to the conceptualization
236 of white noise as a collection of Gaussian distributed random variables. Figure 1 illustrates realizations of
237 this conceptualized white noise in one (left) and two (right) dimensions.

238 However, it turns out that realizations of white noise do not exist as functions in the classical sense. Indeed,
239 since Brownian motion is nowhere differentiable with respect to time, white noise cannot be formally
240 understood as its time derivative. Thus our notation \dot{W} is only meant to aid intuition and not be taken as
241 formal. A formal understanding is possible by considering white noise as a *measure-valued* process (Dawson
242 1975; Walsh 1986) or as a *generalized* process that acts on classically defined functions to return random
243 variables or stochastic processes (Krylov and Rozovskii 1981; Da Prato and Zabczyk 2014). The latter
244 provides the general idea implemented here, though we do not closely follow the treatments of Krylov and
245 Rozovskii (1981) or Da Prato and Zabczyk (2014).

246 Throughout this section, we minimize notation by writing $\int_{\mathbb{R}} f(x)dx = \int_{-\infty}^{+\infty} f(x)dx$ and similarly $\int_D f(x)dx$
247 for the integral of f on $D \subset \mathbb{R}$. We define $L_c^2(\mathbb{R} \times [0, \infty))$ as the set of functions $f(x, t)$ mapping $\mathbb{R} \times [0, \infty)$

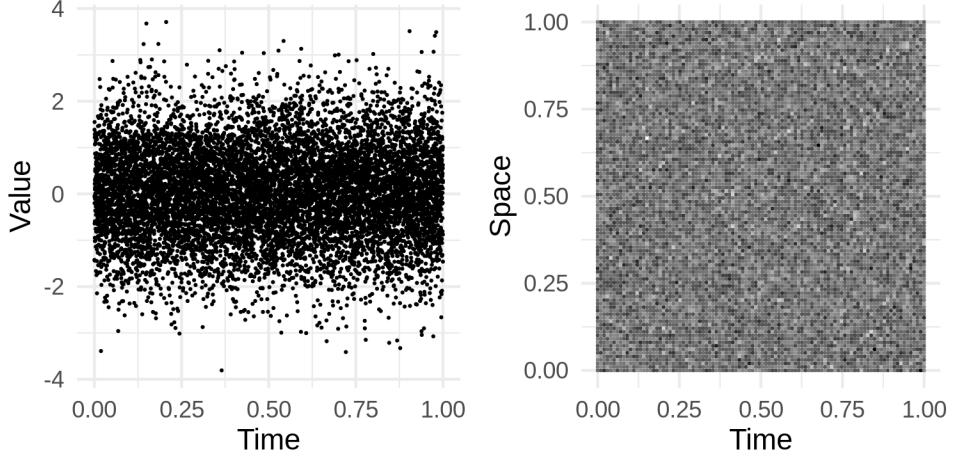


Figure 1: Approximations of sample paths of temporal white noise (left) and space-time white noise (right) with brightness scaled to value.

to \mathbb{R} that are continuous in t and satisfy $\|f\|_2(t) = \sqrt{\int_{\mathbb{R}} f^2(x, t) dx} < \infty$ for each $t \in [0, \infty)$. We define a generalized stochastic process \mathbf{W} that maps functions $f \in L_c^2(\mathbb{R} \times [0, \infty))$ to real-valued stochastic processes indexed by $t \in [0, \infty)$. To evaluate this process for a function $f \in L_c^2(\mathbb{R} \times [0, \infty))$ and some time $t \in [0, \infty)$ we write $\mathbf{W}_t(f)$. Specifically, for any $f, g \in L_c^2(\mathbb{R} \times [0, \infty))$, we define $\mathbf{W}(f)$ and $\mathbf{W}(g)$ to be Gaussian processes satisfying, for any $t, t_1, t_2 \geq 0$,

$$\mathbb{E}(\mathbf{W}_t(f)) = \mathbb{E}(\mathbf{W}_t(g)) = 0, \quad (13)$$

$$\mathbb{C}(\mathbf{W}_{t_1}(f), \mathbf{W}_{t_2}(g)) = \int_0^{t_1 \wedge t_2} \int_{\mathbb{R}} f(x, s) g(x, s) dx ds, \quad (14)$$

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

Since Gaussian processes are characterized by their expectations and covariances, the processes $\mathbf{W}(f)$ and $\mathbf{W}(g)$ are well defined (under the convention $f = g$ whenever $\|f - g\|_2(t) = 0$ for each $t \geq 0$). In particular, if $f \in L_c^2(\mathbb{R} \times [0, \infty))$ is independent of time, then $\mathbf{W}(f)$ is a Brownian motion with variance at time $t \geq 0$ equal to $t\|f\|_2^2$. With the generalized process \mathbf{W} defined, we define the space-time white noise $\dot{W}(x, t)$ implicitly via the stochastic integral

$$\left\langle \int_0^t \int_{\mathbb{R}} f(x, s) \dot{W}(x, s) dx ds \right\rangle = \left\langle \int_{\mathbb{R}} \int_0^t f(x, s) \dot{W}(x, s) dx ds \right\rangle = \mathbf{W}_t(f), \quad \forall f \in L_c^2(\mathbb{R} \times [0, \infty)), \quad t \geq 0. \quad (15)$$

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

$$I_{D \times [0, \infty)}(x, t) = \begin{cases} 0, & x \notin D \\ 1, & x \in D \end{cases}. \quad (16)$$

Since $\int_{\mathbb{R}} I_{D \times [0, \infty)}^2(x, t) dx = \int_D dx = |D|$, where $|D|$ denotes the length of D , $I_{D \times [0, \infty)} \in L_c^2(\mathbb{R} \times [0, \infty))$. Thus, by equations (13) and (14), we have

$$\mathbb{E} \left(\int_0^t \int_D \dot{W}(x, s) dx ds \right) = 0, \quad (17)$$

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

269 where \mathbb{V} denotes the variance operator with respect to the underlying probability space. Following this
270 notation, equations (13) and (14) can be rewritten as

$$\mathbb{E} \left(\int_0^t \int_{\mathbb{R}} f(x, s) \dot{W}(x, s) dx ds \right) = 0, \quad (19)$$

271

$$\mathbb{C} \left(\int_0^{t_1} \int_{\mathbb{R}} f(x, s) \dot{W}(x, s) dx ds, \int_0^{t_2} \int_{\mathbb{R}} g(x, s) \dot{W}(x, s) dx ds \right) = \int_0^{t_1 \wedge t_2} \int_{\mathbb{R}} f(x, s) g(x, s) dx ds. \quad (20)$$

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

$$\int_0^t d\hat{\mathbf{W}}_s(f) = \int_0^t \int_{\mathbb{R}} \frac{f(x, s)}{\sqrt{\int_{\mathbb{R}} f^2(s, y) dy}} \dot{W}(x, s) dx ds. \quad (21)$$

274 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 , $\int_0^t d\hat{\mathbf{W}}_s(f)$
275 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
276 shorthand used to denote stochastic differentials. Thus, equation (20) effectively extends Itô's multiplication
277 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 |

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

282 2.2.2 From branching processes to SPDE

283 In real populations individuals are born and potentially reproduce before they ultimately die. These three
284 events provide the basic ingredients of a branching process. Mathematical investigations of such processes have
285 a relatively deep history (Kendall 1966). The most simple branching process, known as the Galton-Watson
286 process, describes the number of individuals alive at a given time $t \geq 0$ as a non-negative integer (Kimmel and
287 Axelrod 2015). Feller (1951) introduced a formal method to approximate branching processes with diffusion
288 processes which are continuous in state (i.e., population size is approximated as a continuous quantity). Since
289 diffusion processes possess greater analytical tractability than branching processes, Feller's method, known as
290 the diffusion limit, has acquired immense popularity particularly in the field of mathematical population
291 genetics (Ewens 2004). For over the past half of a century a great deal of accomplishments have been achieved
292 in formalizing the diffusion limits of branching processes that describe populations of individuals occurring in
293 some continuous space (Watanabe 1968; Dawson 1975; Perkins 1992, 1995; Méléard and Roelly 1993; Li 1998;
294 Bertoin and Le Gall 2003; Etheridge 2008; Barton and Etheridge 2019). This space can represent geographic

space or, relevant to our context, phenotypic space. In the following subsection, we describe the BBM process, which is a particularly important branching process with spatial structure. This process has been very useful in the study of SPDE due to its simplifying assumption that individuals do not interact. However, this assumption imposes an unfortunate restriction by precluding the modelling of ecological interactions. We therefore follow our discussion of BBM with a review of a few important results on spatially structured branching processes that account for interactions.

Branching Brownian motion

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

- 1) **Branching rate:** Lifetimes of individuals are assumed to be exponentially distributed with death rate $\lambda > 0$. Reproduction is assumed to occur simultaneously with death.
- 2) **Reproductive law:** At the death of an individual, a branching event occurs where a random (possibly zero) number of offspring are left. The distribution of offspring left is called the reproductive law or branching mechanism. We denote the mean and variance in reproductive output by \mathcal{W} and V respectively. The case of $\mathcal{W} = 1$ is referred to as the critical condition.
- 3) **Spatial movement:** Each offspring is born at the current location of their parent. Immediately after birth they move around space according to Brownian motion with diffusion parameter $\sqrt{\mu}$. In our context we interpret spatial movement as mutation so that the location of an individual at death represents the value of its phenotype. Then an individual born at location $x \in \mathbb{R}^d$ that lives for $\tau > 0$ units of time will have a normally distributed trait centered on x with covariance matrix equal to $\tau\mu$ times the $d \times d$ identity matrix. Hence, offspring inherit normally distributed traits centered on their parental trait. This fact creates a vital link to the deterministic dynamics reviewed above. Indeed, in the absence of selection, the deterministic PDE (7) reduces to the $d = 1$ -dimensional Kolmogorov forward equation for a Brownian motion with diffusion parameter $\sqrt{\mu}$.

To obtain a SPDE from a BBM we take a diffusion limit. There are several ways to do this, but a simple approach is to rescale the mass of individuals and time by $1/n$, diffusion by $\mu \rightarrow \mu/n$, branching rate by $\lambda \rightarrow n\lambda$, fitness by $\mathcal{W} \rightarrow \mathcal{W}^{1/n}$ and consider the limit as $n \rightarrow \infty$. Denoting the rescaled process by $X_t^{(n)}(D)$, 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 measure-valued process (Watanabe 1968). Instead of returning the number of individuals alive in a region of space, super-Brownian motion returns the *mass* of the population concentrated in a region of space. In particular, the value of $\mathcal{X}_t(D)$ is generally a continuously varying non-negative random variable for any $t \geq 0$ and $D \subset \mathbb{R}^d$. It turns out that for spatial dimension $d = 1$, \mathcal{X}_t is absolutely continuous with respect to the Lebesgue measure for each $t \geq 0$ (Konno and Shiga 1988; Reimers 1989). This means that we can write $\mathcal{X}_t(D) = \int_D \nu(x, t) dx$ for some density process $\nu(x, t)$. Setting $\lambda = 1$ and $m = \ln \mathcal{W}$ this density process satisfies the SPDE

$$\frac{\partial}{\partial t} \nu(x, t) = m\nu(x, t) + \frac{\mu}{2} \frac{\partial^2}{\partial x^2} \nu(x, t) + \sqrt{V\nu(x, t)} \dot{W}(x, t). \quad (22)$$

Since $\nu(x, t)$ is not generally differentiable in x or t , the derivatives in expression (22) are taken in the weak sense. That is, to rigorously interpret SPDE (22), we integrate the solution $\nu(x, t)$ against functions $f \in C_b^2(\mathbb{R})$, where $C_b^2(\mathbb{R})$ is the set of bounded and twice continuously differentiable functions on \mathbb{R} . Hence, equation (22) is just an abbreviation for

$$\begin{aligned} & \int_{\mathbb{R}} \nu(x, t) f(x) dx - \int_{\mathbb{R}} \nu(x, 0) f(x) dx \\ &= \int_0^t \int_{\mathbb{R}} \nu(x, s) \left(m f(x) + \frac{\mu}{2} \frac{\partial^2}{\partial x^2} f(x) \right) ds dx + \int_0^t \int_{\mathbb{R}} f(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) dx ds. \quad (23) \end{aligned}$$

339 For more on the general theory of SPDE see Walsh (1986). Note that since $\nu(x, t)$ is the density of a finite
 340 measure, it is integrable for each $t \geq 0$. Thus, since for some $M > 0$, $|f(x)| \leq M$ for every $x \in \mathbb{R}$, setting
 341 $\varphi(x, t) = f(x) \sqrt{V \nu(x, t)}$ implies $\varphi \in L_c^2(\mathbb{R} \times [0, \infty))$. Hence, the white noise integral on the right-hand side
 342 of equation (23) can be understood using the heuristics introduced above. Evaluating equation (23) in the
 343 particular case of $f(x) \equiv 1$ returns the total mass process, which we refer to as the total abundance $N(t)$.

344 A convergence theorem for the diffusion limit of a generalization of BBM was established by Watanabe
 345 (1968). Dawson (1975) suggested that, for spatial dimension $d = 1$, this diffusion limit should admit a density
 346 process that satisfies a SPDE. Konna and Shiga (1988) and Reimers (1989) independently proved Dawson's
 347 suggestion was indeed correct. The diffusion limit of this more general branching process (in arbitrary
 348 spatial dimension) is referred to as the Dawson-Watanabe superprocess (Etheridge 2000). Conditioning the
 349 Dawson-Watanabe superprocess to have constant mass returns the Fleming-Viot process (Etheridge and
 350 March 1991; Perkins 1991) which has been popular in studies of spatial population genetics. In particular, an
 351 extension of the Fleming-Viot process, known as the Λ -Fleming-Viot process, was introduced by Bertoin
 352 and Le Gall (2003) and coined by Etheridge (2008) where it was used to resolve some technical challenges in
 353 modelling isolation by distance (Felsenstein 1975; see also Barton, Etheridge, and Véber 2013; and Barton
 354 and Etheridge 2019). Although this provides an impressive list of accomplishments, the Dawson-Watanabe
 355 superprocess falls short of our needs. In particular this process assumes individuals do not interact and thus
 356 precludes its ability to model ecological interactions. However, this concern has been addressed, leading to
 357 constructions of superprocesses that account for interactions among individuals. In the next subsection we
 358 summarize the main results in this area and introduce the SPDE that provides the basis for our approach to
 359 theoretical evolutionary ecology.

360 Interacting superprocesses

361 The existence of diffusion limits for a class of measure-valued branching processes involving interactions
 362 among individuals has been treated by Méléard and Roelly (1992, 1993). The interactions can manifest as
 363 dependencies of the spatial movement or reproductive law of individuals on their position x and the state of
 364 the whole population described by X_t . An important result of Méléard and Roelly (1992, 1993) is a theorem
 365 that provides sufficient conditions to construct a sequence of rescaled measure-valued branching processes
 366 that converge to a generalization of the Dawson-Watanabe superprocess that includes interactions. The
 367 rescaling is analogous to that described above for non-interacting Dawson-Watanabe superprocesses, but now
 368 the reproductive law described by $\mathcal{W}(X_t, x)$ and $V(X_t, x)$, branching rate $\lambda(X_t, x)$ and diffusion parameter
 369 $\sqrt{\mu(X_t, x)}$ are allowed to depend on the whole population X_t and individual location x . In Figure 2 we
 370 demonstrate this rescaling in discrete time for a population experiencing stabilizing selection and logistic
 371 growth. Since time is discretized, the process we simulate is formally a branching random walk. For further
 372 details on our simulation see SM §5.4.

373 Interactions that manifest in the spatial movement can be used to model mutation bias and those manifesting
 374 in the reproductive law can model density-dependent growth rates and frequency-dependent selection. Perkins
 375 (1992, 1995) developed a theory of stochastic integration with respect to the so-called *Brownian trees* to
 376 characterize interacting superprocesses and establish properties of existence and uniqueness. Li (1998) built
 377 directly off of the construction of Méléard and Roelly (1992, 1993) to study properties of density processes
 378 associated with interacting superprocesses, arriving at a SPDE that forms the foundation of our approach.

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

Rescalings of a branching random walk

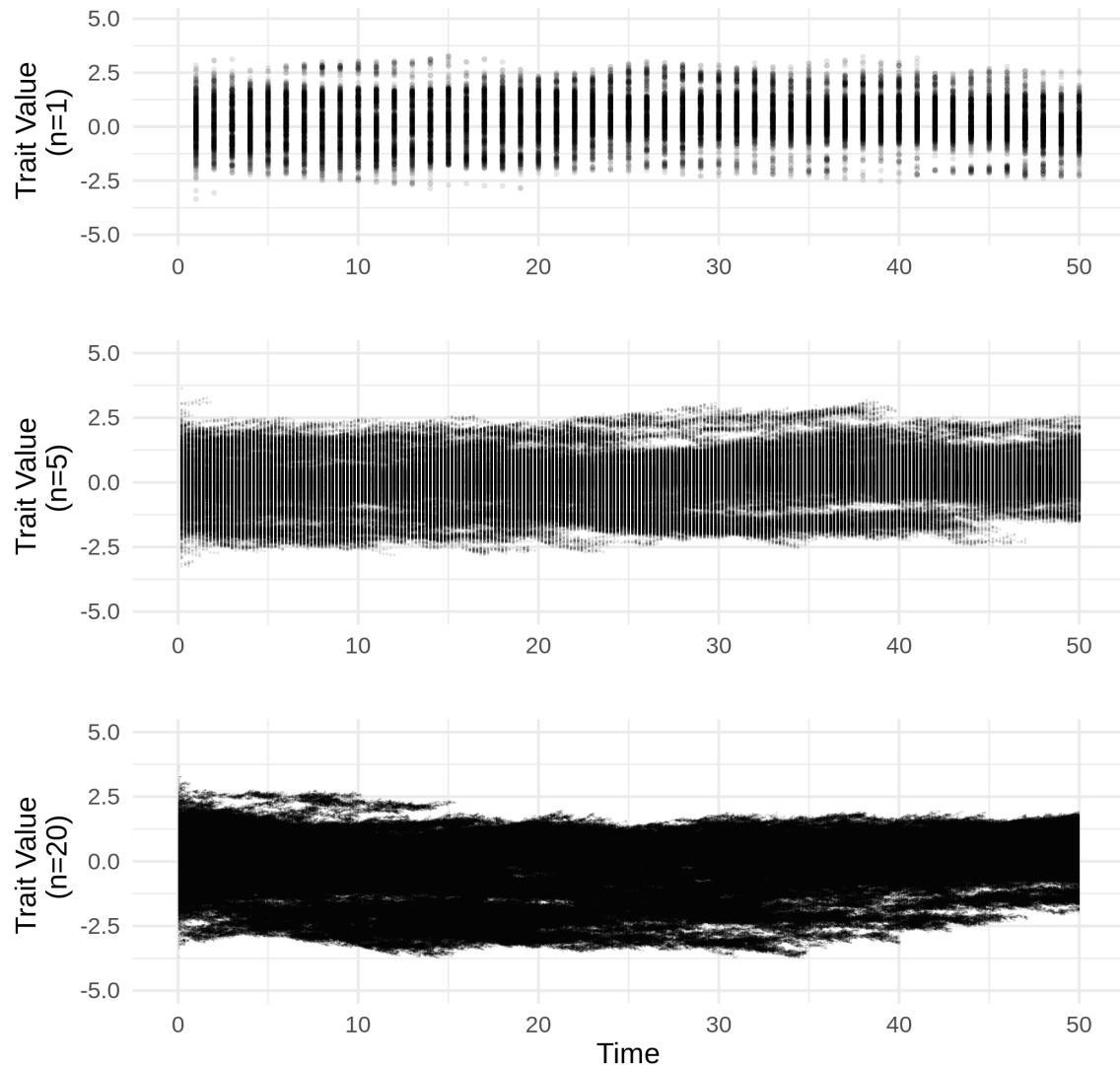


Figure 2: Rescaled sample paths of a branching random walk under stabilizing selection and logistic growth. The top plot displays a sample path without scaling ($n = 1$), the middle plot shows a sample path rescaled by $n = 5$ and the bottom plot shows a sample path rescaled by $n = 20$.

383 space and satisfies the SPDE

$$\frac{\partial}{\partial t}\nu(x,t) = m(\nu,x)\nu(x,t) + \frac{\mu}{2}\frac{\partial^2}{\partial x^2}\nu(x,t) + \sqrt{V\nu(x,t)}\dot{W}(x,t). \quad (24)$$

384 Note that, comparing equation (24) to equation (3.5) of Li (1998), our m and V correspond to Li's b and c
385 respectively. Denoting $C_1^+(\mathbb{R})$ the set of non-negative, continuous, integrable functions on \mathbb{R} , it is important
386 to note that, under the assumptions made in Méléard and Roelly (1992, 1993) and Li (1998), equation (24) is
387 only formal when $m(\nu,x)$ is bounded across all combinations of $\nu \in C_1^+(\mathbb{R})$ and $x \in \mathbb{R}$. However, recalling
388 our condition $m(\nu,x) \leq r \in \mathbb{R}$, the growth rates we consider are only bounded above. Replacing $m(\nu,x)$ with
389 its upper-bound demonstrates that the associated superprocess should still admit an integrable density. That
390 is, we should still be able to compute the total mass process $N(t)$. In fact, in the proof of the construction of
391 the interacting superprocess as the limit of rescaled branching diffusions, Méléard and Roelly (1992, 1993)
392 assumed $m(\nu,x)$ to be bounded to guarantee the total mass process will have finite mean and variance. This
393 allowed them to employ a tightness criterion for sequences of measures and show the rescaled processes
394 converge to a superprocess with finite total mass. Li's (1998) result builds directly off of Méléard and Roelly's
395 construction, inheriting the assumption of boundedness for $m(\nu,x)$. However, in Li (1998), the sufficiency of
396 $m(\nu,x)$ being bounded above is even more clear since Li works explicitly with a common upperbound for
397 both $m(\nu,x)$ and V . Hence, one can repeat the necessary proofs replacing the assumption that $m(\nu,x)$ is
398 bounded with the assumption that $m(\nu,x)$ is merely bounded above to derive the same results.

399 What remains is to show the spatial mean and variance are finite, assuming finite initial conditions. That is,
400 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
401 each $t > 0$ (it turns out we need $M_4(t) < \infty$ for the SDE describing the evolution of phenotypic variance). By
402 condition (??), we see that $m(\nu,x)$ drops quadratically as x leaves θ . In SM §5.1 we show that, when $V = 0$,
403 this form of fitness generates a restoring force that ensures a finite mean and variance. Since mutation causes
404 finite perturbations we do not expect it to overcome this restoring force in the stochastic case. If anything,
405 we suspect a large mutation variance to merely increase the genetic load of the population and drive it to
406 extinction. Indeed, this suggestion is supported by equilibrium of the corresponding deterministic system
407 found in SM §5.4. However, to establish further support, in SM §5.4 we investigate these claims by comparing
408 simulations of the BBM with selection to analytical solutions of the PDE obtained from equation (24) by
409 setting $V = 0$. The results suggest our assertions are sound, but they have yet to be rigorously justified. For
410 now, we leave this as an open problem and assume spatial means and variances are finite for every $t \in [0, \infty)$.
411 In the next subsection we derive an SDE for the abundance of a population and report the SDE for the trait
412 mean \bar{x} and trait variance σ^2 . For the sake of space, derivations of $d\bar{x}$ and $d\sigma^2$ can be found in SM §5.5.

413 2.2.3 Deriving SDE from a SPDE

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

$$\begin{aligned} & \int_{\mathbb{R}} f(x)\nu(x,t)dx - \int_{\mathbb{R}} f(x)\nu(x,0)dx \\ &= \int_0^t \int_{\mathbb{R}} \nu(x,s) \left(m(\nu,x)f(x) + \frac{\mu}{2}\frac{\partial^2}{\partial x^2}f(x) \right) dxds + \int_0^t \int_{\mathbb{R}} f(x)\sqrt{V\nu(x,s)}\dot{W}(x,s)dxds \end{aligned} \quad (25)$$

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

$$N(t) - N(0) = \bar{m}(t)N(t) + \int_0^t \int_{\mathbb{R}} \sqrt{V\nu(x,s)}\dot{W}(x,s)dxds. \quad (26)$$

417 Equation (20) implies

$$\mathbb{V} \left(\int_0^t \int_{\mathbb{R}} \sqrt{V\nu(x,s)}\dot{W}(x,s)dxds \right) = \int_0^t \int_{\mathbb{R}} V\nu(x,s)dxds = \int_0^t VN(s)ds. \quad (27)$$

⁴¹⁸ Following the notation introduced in §2.2.1, we have

$$\int_0^t d\hat{\mathbf{W}}_s(\sqrt{V\nu(x,s)}) = \int_0^t \int_{\mathbb{R}} \frac{\sqrt{V\nu(x,s)}}{\sqrt{VN(s)}} \dot{W}(x,s) dx ds \quad (28)$$

⁴¹⁹ is, as a function of t , a standard Brownian motion. To clean up notation set $dW_1(t) = d\hat{\mathbf{W}}_t(\sqrt{V\nu(x,t)})$.
⁴²⁰ Then we can write an SDE for N which we provide in the following section.

⁴²¹ Unfortunately, the same trick does not work for calculating SDE for the mean and variance of traits. Indeed,
⁴²² setting $f(x) = x, x^2$ and evaluating expression (25) will lead to SDE that miscalculate the effect of mutation
⁴²³ and hence are clearly incorrect results. Obtaining the correct results turns out to require some very involved
⁴²⁴ calculations which we perform in SM §5.5.

⁴²⁵ 2.3 Equations of evolutionary and demographic dynamics

⁴²⁶ In SM §5.5 we use our assumptions that trait mean and variance are finite and that variance in reproductive
⁴²⁷ 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

$$dN(t) = \bar{m}(t)N(t)dt + \sqrt{VN(t)}dW_1(t), \quad (29a)$$

$$d\bar{x}(t) = \text{Cov}_t(x, m(\nu, x))dt + \frac{\mu}{2} \int_{-\infty}^{+\infty} x \frac{\partial^2}{\partial x^2} p(x, t)dx + \sqrt{V \frac{\sigma^2(t)}{N(t)}} dW_2(t), \quad (29b)$$

$$d\sigma^2(t) = \left(\text{Cov}_t((x - \bar{x}(t))^2, m(\nu, x)) + \frac{\mu}{2} \int_{-\infty}^{+\infty} (x - \bar{x}(t))^2 \frac{\partial^2}{\partial x^2} p(x, t)dx - V \frac{\sigma^2(t)}{N(t)} \right) dt \\ + \sqrt{V \frac{(x - \bar{x}(t))^4 - \sigma^4(t)}{N(t)}} dW_3(t). \quad (29c)$$

⁴²⁹ where W_1 , W_2 and W_3 are standard Brownian motions. Dividing by dt one can interpret equations (29) as if
⁴³⁰ they are ordinary differential equations, but this not technically rigorous since Brownian motion is nowhere
⁴³¹ differentiable with respect to time. In SM §5.5 we show that in general W_1 is independent of both W_2 and
⁴³² W_3 , but W_2 and W_3 covary.

⁴³³ There is quite a bit we can learn from expressions (29). Firstly, setting $V = 0$ recovers the deterministic
⁴³⁴ dynamics derived in §2.1. Alternatively, one can take $N(t) \rightarrow \infty$ to recover the deterministic dynamics for
⁴³⁵ $\bar{x}(t)$ and $\sigma^2(t)$. Characteristically, we note the effect of demographic stochasticity on abundance grows with
⁴³⁶ $\sqrt{N(t)}$. Hence, dividing by N , we find the effects of demographic stochasticity on the per-capita growth rate
⁴³⁷ diminish with increased abundance. Relating the response to demographic stochasticity derived here to the
⁴³⁸ effect of random genetic drift derived in classic quantitative genetic theory, if we set $\sigma^2(t) = \sigma^2$ and $N(t) = N$
⁴³⁹ constant with respect to time, then integrating the stochastic term in equation (29b) over a single unit of time
⁴⁴⁰ returns a normally distributed random variable with mean zero and variance equal to $V\sigma^2/N$. In particular,
⁴⁴¹ assuming perfect inheritance, when reproductive variance is unity ($V = 1$) this random variable coincides
⁴⁴² with the effect of random genetic drift on the change in mean trait over a single generation derived using
⁴⁴³ sampling arguments (Lande 1976). There is also an interesting connection with classical population genetics.
⁴⁴⁴ A fundamental result from early population genetic theory is the expected reduction in diversity due to the
⁴⁴⁵ chance loss of alleles in finite populations (Fisher 1923; Wright 1931). This expected reduction in diversity
⁴⁴⁶ due to random genetic drift is captured by the third term in the deterministic component of expression
⁴⁴⁷ (29c), particularly $-V\sigma^2(t)/N(t)$. The component of SDE (29c) describing random fluctuations in $\sigma^2(t)$ is
⁴⁴⁸ more complicated and is proportional to the root of the difference between the centralized fourth moment of
⁴⁴⁹ $p(x, t)$ and $\sigma^4(t)$. These expressions can be used to investigate the dynamics of the mean and variance for
⁴⁵⁰ general $\nu(x, t)$. However, in the next subsection we simplify these expressions by approximating $\nu(x, t)$ with
⁴⁵¹ a Gaussian curve. In SM §5.5 we show that under the Gaussian case W_1 , W_2 and W_3 are independent.

452 **2.3.1 Particular results assuming a Gaussian phenotypic distribution**

453 By assuming $\nu(x, t)$ can be approximated by a Gaussian curve for each $t \geq 0$, expressions (29a), (29b)
 454 and (29b) transform into efficient tools for deriving the dynamics of populations given a fitness function
 455 $m(\nu, x)$. Gaussian phenotypic distributions are often obtained through Gaussian, exponential or weak selection
 456 approximations together with a simplified model of inheritance and random mating (Lande 1980; Turelli
 457 1984, 1986, 2017; Bürger 2000). Alternatively, it has been shown that a Gaussian distribution can provide a
 458 reasonable approximation even when selection is strong and non-Gaussian (Turelli and Barton 1994). However,
 459 our approach adds an additional layer of difficulty. Even with Gaussian selection, the resulting solution to
 460 SPDE (24) will only be a Gaussian curve in expectation, assuming a Gaussian initial condition. Yet this
 461 difficulty is not as challenging as it may first appear. Indeed, since SPDE (24) can be derived as a diffusion
 462 limit we know that, under the appropriate assumptions on selection, genetic architecture and reproduction,
 463 the stochastic departure from a Gaussian curve is negligible when the ratio V/N is small (i.e., when the
 464 variance in reproductive output is much smaller than the population size). In SM §5.4 we demonstrate this
 465 result using numerical methods. Mathematically, this requirement restricts model parameters to regions
 466 that maintain large population sizes. Biologically, this implies populations are not at risk of extinction.
 467 Hence, models developed in this framework are not suitable for studying colonization-extinction dynamics or
 468 evolutionary rescue. Allowing for these restrictions, we may safely assume that ν is approximately Gaussian
 469 and justify writing

$$\nu(x, t) = \frac{N(t)}{\sqrt{2\pi\sigma^2(t)}} \exp\left(-\frac{(x - \bar{x}(t))^2}{2\sigma^2(t)}\right). \quad (30)$$

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

$$\text{Cov}(x, m) = \sigma^2 \left(\frac{\partial \bar{m}}{\partial \bar{x}} - \overline{\frac{\partial m}{\partial x}} \right). \quad (31)$$

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

$$\text{Cov}\left((x - \bar{x})^2, m\right) = 2\sigma^4 \left(\frac{\partial \bar{m}}{\partial \sigma^2} - \overline{\frac{\partial m}{\partial \sigma^2}} \right) \quad (32)$$

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

$$d\bar{x} = \sigma^2 \left(\frac{\partial \bar{m}}{\partial \bar{x}} - \overline{\frac{\partial m}{\partial x}} \right) dt + \sqrt{V \frac{\sigma^2}{N}} dW_2, \quad (33a)$$

$$d\sigma^2 = 2\sigma^4 \left(\frac{\partial \bar{m}}{\partial \sigma^2} - \overline{\frac{\partial m}{\partial \sigma^2}} \right) dt + \left(\mu - V \frac{\sigma^2}{N} \right) dt + \sigma^2 \sqrt{\frac{2V}{N}} dW_3. \quad (33b)$$

473 474 These equations allow us to derive the response in trait mean and variance by taking derivatives of fitness, a
 475 much more straightforward operation than calculating a covariance for general phenotypic distributions. Note
 476 that in the above expressions, the partial derivatives of \bar{m} represent frequency independent selection and the
 477 averaged partial derivatives of m represent frequency dependent selection. This relationship has already been
 478 pointed out by Lande (1976) for the evolution of the mean trait, but here we see an analogous relationship
 479 holds also for the evolution of trait variance.

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

485 **2.3.2 The evolution of additive genetic variance**

486 To model imperfect heritability we consider the relationship between expressed phenotypes $x \in \mathbb{R}$ and
 487 associated genetic values $g \in \mathbb{R}$ known as *breeding values*. The breeding value of an individual is the sum
 488 of additive effects of the alleles carried by the individual on its expressed trait. Since our derivations of
 489 evolutionary equations are based on branching processes that assume asexually reproducing populations
 490 (§2.2.2), the additive genetic variance G is just the variance of breeding values in a population. For a detailed
 491 treatment of breeding values and additive genetic variances, see Bulmer (1980) and Lynch and Walsh (1998).
 492 Following standard assumptions of classical quantitative genetics we assume that the expressed trait for any
 493 given individual is normally distributed around their breeding value with variance η . Hence, $\sigma^2 = G + \eta$.
 494 This coincides with assuming breeding values can be predicted from expressed traits using ordinary least
 495 squares. In the case that all of the effects of alleles on an expressed trait are additive, η is known as the
 496 *variance of environmental deviation* (Lynch and Walsh 1998). For a given breeding value, we denote the
 497 probability density of a randomly drawn expressed trait by $\psi(x, g)$. Hence,

$$\psi(x, g) = \frac{1}{\sqrt{2\pi\eta}} \exp\left(-\frac{(x-g)^2}{2\eta}\right). \quad (34)$$

498 To include this relationship in our framework, we write $\rho(g, t)$ as the abundance density of breeding values
 499 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
 500 evolution of $\nu(x, t)$ to modelling the evolution of $\rho(g, t)$. Once $\rho(g, t)$ is determined, we can compute $\nu(x, t)$
 501 via

$$\nu(x, t) = \int_{-\infty}^{+\infty} \rho(g, t) \psi(x, g) dg. \quad (35)$$

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

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

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

$$\dot{\rho}(g, t) = m^*(\rho, g) \rho(g, t) + \frac{\mu}{2} \frac{\partial^2}{\partial^2 g} \rho(g, t) + \sqrt{V\rho(g, t)} \dot{W}(g, t). \quad (37)$$

515 We assume $\rho(g, t)$ is Gaussian which implies its mode coincides with \bar{x} . Furthermore, since $\sigma^2 = G + \eta$, we
 516 can use equation (36) and the chain rule from calculus to find

$$\frac{\partial \bar{m}}{\partial G} = \frac{\partial \bar{m}}{\partial \sigma^2} \frac{\partial \sigma^2}{\partial G} = \frac{\partial \bar{m}}{\partial \sigma^2}, \quad (38a)$$

$$\frac{\partial \bar{m}}{\partial G} = \frac{\partial \bar{m}}{\partial \sigma^2} \frac{\partial \sigma^2}{\partial G} = \frac{\partial \bar{m}}{\partial \sigma^2}. \quad (38b)$$

518 Thus, equations (33) become

519

$$d\bar{x} = G \left(\frac{\partial \bar{m}}{\partial \bar{x}} - \frac{\overline{\partial m}}{\partial \bar{x}} \right) dt + \sqrt{V \frac{G}{N}} dW_2, \quad (39a)$$

$$dG = 2G^2 \left(\frac{\partial \bar{m}}{\partial G} - \frac{\overline{\partial m}}{\partial G} \right) dt + \left(\mu - V \frac{G}{N} \right) dt + G \sqrt{\frac{2V}{N}} dW_3. \quad (39b)$$

520 From expressions (39) we see that, under our model of inheritance, focusing on additive genetic variance G
 521 instead the variance in expressed traits σ^2 makes no structural changes to the basic equations describing the
 522 dynamics of populations.

523 3 A model of diffuse coevolution

524 3.1 Formulation

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

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

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

$$u_i(\zeta, x_i) = \frac{U_i}{\sqrt{2\pi w_i}} \exp \left(-\frac{(x_i - \zeta)^2}{2w_i} \right). \quad (40)$$

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

$$e_i(\zeta) = Q_i \exp \left(-\frac{A_i}{2} (\theta_i - \zeta)^2 \right). \quad (41)$$

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

$$\int_{-\infty}^{+\infty} e_i(\zeta) u_i(\zeta, x_i) d\zeta = \frac{Q_i U_i}{\sqrt{A_i w_i + 1}} \exp \left(-\frac{A_i}{2(A_i w_i + 1)} (\theta_i - x_i)^2 \right). \quad (42)$$

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

$$\mathcal{O}_{ij}(x_i, x_j) = \int_{-\infty}^{+\infty} u_i(\zeta, x_i) u_j(\zeta, x_j) d\zeta = \frac{U_i U_j}{\sqrt{2\pi(w_i + w_j)}} \exp\left(-\frac{(x_i - x_j)^2}{2(w_i + w_j)}\right). \quad (43)$$

550 A notable criticism of using niche overlap to measure the intensity of competition points to cases where
 551 populations competing on multiple niche axes exhibit overlap on at least one of the axes, but no overall
 552 niche overlap (Holt 1987). Thus niche overlap on lower-dimensional projections of some multivariate niche
 553 space does not imply the populations compete. To illustrate with a simple example, consider two populations
 554 competing for space on the plane \mathbb{R}^2 . If the spatial distributions of the two populations overlap, then
 555 they will overlap on both spatial axes. However, if the populations do not overlap on at least one of the
 556 spatial axes, they will have no overall spatial overlap. Furthermore, even if the species overlap on both
 557 spatial axes, they need not have any overall spatial overlap. This final result corresponds to the fact that
 558 components of niche space do not necessarily interact multiplicatively to determine the consequences for
 559 the intensity of competition. In another component of Holt's (1987) critique, an argument is made for the
 560 potential of competition occurring without any overlap in niche space. However, this argument is based on
 561 the practical difficulty of identifying every resource axis populations are competing on and how these axes
 562 interact to determine fitness consequences. Our model avoids these caveats by assuming competition only
 563 occurs along a single dimensional resource gradient. To map the degree of niche overlap to fitness, we assume
 564 competition between individuals with niche locations x_i and x_j additively decreases the Malthusian fitness
 565 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
 566 equations (29a), (39a) and (39b) to find

$$dN_i = \left\{ R_i - \frac{a_i}{2} ((\bar{x}_i - \theta_i)^2 + G_i + \eta_i) - c_i \sum_{j=1}^S N_j U_i U_j \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2} (\bar{x}_i - \bar{x}_j)^2} \right\} N_i dt + \sqrt{V_i N_i} dW_1, \quad (44a)$$

$$d\bar{x}_i = \left\{ a_i G_i (\theta_i - \bar{x}_i) - c_i G_i \left(\sum_{j=1}^S N_j U_i U_j b_{ij} (\bar{x}_j - \bar{x}_i) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2} (\bar{x}_i - \bar{x}_j)^2} \right) \right\} dt + \sqrt{V_i \frac{G_i}{N_i}} dW_2, \quad (44b)$$

$$dG_i = \left\{ c_i G_i^2 \left(N_i U_i^2 b_{ii} \sqrt{\frac{b_{ii}}{2\pi}} + \sum_{j=1}^S N_j U_i U_j b_{ij} (1 - b_{ij} (\bar{x}_i - \bar{x}_j)^2) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2} (\bar{x}_i - \bar{x}_j)^2} \right) \right. \\ \left. + \mu_i - a_i G_i^2 - V_i \frac{G_i}{N_i} \right\} dt + G_i \sqrt{\frac{2V_i}{N_i}} dW_3, \quad (44c)$$

567 where

$$R_i = \ln \left(\frac{Q_i U_i}{\sqrt{1 + A_i w_i}} \right), \quad (45a)$$

$$a_i = \frac{A_i}{1 + A_i w_i}, \quad (45b)$$

$$b_{ij}(t) = b_{ji}(t) = (w_i + w_j + \eta_i + \eta_j + G_i(t) + G_j(t))^{-1}, \quad (45c)$$

$$c_i \geq 0. \quad (45d)$$

568 Despite the convoluted appearance of system (44), there are some nice features that reflect biological reasoning.
 569 For example, the dynamics of abundance are just a generalization of Lotka-Volterra dynamics. In particular,

570 the effect of competition with species j on the fitness of species i grows linearly with N_j . However, as biotic
 571 selection pushes \bar{x}_i away from \bar{x}_j , the effect of competition with species j on the fitness of species i rapidly
 572 diminishes, reflecting a reduction in niche overlap. The divergence of \bar{x}_i and \bar{x}_j due to competition is referred
 573 to in the community ecology literature as character displacement. We also see that the fitness of species i
 574 drops quadratically with the difference between \bar{x}_i and the abiotic optimum θ_i . Hence, abiotic selection acts
 575 to pull \bar{x}_i towards θ_i . The response in mean trait \bar{x}_i to natural selection is proportional to the amount of
 576 heritable variation in the population, represented by the additive genetic variance G_i . However, we have
 577 that G_i is itself a dynamic quantity. Under our model, abiotic stabilizing selection erodes away heritable
 578 variation at a rate that is independent of both N_i and \bar{x}_i . The effect of competition on G_i is a bit more
 579 complicated. When $b_{ij}(\bar{x}_i - \bar{x}_j)^2 < 1$, competition with species j acts as diversifying selection which tends to
 580 increase the amount of heritable variation. However, when $b_{ij}(\bar{x}_i - \bar{x}_j)^2 > 1$, competition with species j acts
 581 as directional selection and reduces G_i . In the following subsections we demonstrate the behavior of system
 582 (44) by plotting numerical solutions and investigate implications for the relationship between the strength of
 583 ecological interactions and selection.

584 3.2 Community dynamics

585 For the sake of illustration we numerically integrated system (44) for a richness of $S = 100$ species. We
 586 assumed homogeneous model parameters across species in the community as summarized by Table 2. We
 587 repeated numerical integration under the two scenarios of weak and strong competition. For the first
 588 scenario of weak competition we set $c = 1.0 \times 10^{-7}$ and for the second scenario of strong competition we set
 589 $c = 5.0 \times 10^{-6}$. With these two sets of model parameters, we simulated our model for 1000.0 units of time.
 590 For both scenarios, we initialized the trait means to $\bar{x}_i = 0.0$, additive genetic variances to $G_i = 10.0$ and
 591 abundances to $N_i = 1000.0$ for each $i = 1, \dots, S$.

592 Temporal dynamics for each scenario are provided in Figure 3. This figure suggests weaker competition leads
 593 to smoother dynamics and a higher degree of organization within the community. Considering expression
 594 (44a) we note that, all else equal, relaxed competition allows for larger growth rates which promote greater
 595 abundances. From (44a) we also note that the per-capita effects on demographic stochasticity diminish with
 596 abundance. To see this, divide both sides by N_i . Inspecting expressions (44b) and (44c), we see that larger
 597 abundances also erode the effects of demographic stochasticity on the evolution of mean trait and additive
 598 genetic variance. These effects were already noted in §2.3, and thus are not a consequence of our model of
 599 coevolution per-se, but we revisit them here since Figure 3 demonstrates the importance of demographic
 600 stochasticity in structuring ecological communities even when populations are very large. Hence, contrary to
 601 the common assumption that stochastic effects can be ignored for large populations, we find that minute
 602 asymmetries by demographic stochasticity remain significant drivers of community structure. In particular,
 603 although we initialized the species with identical state variables and model parameters, we find an enormous
 604 amount of asymmetry and even some potential phase changes. In the following two paragraphs we describe
 605 the natural history of the community as illustrated in Figure 3.

606 After about 125.0 units of time, the community appears to have shaken off the initial conditions and entered
 607 into a qualitatively distinct phase of dynamics. Aside from a few outliers, most of the species remain clustered
 608 together in their state variables. This lasts for approximately 375.0 units of time until, at around time 500.0,
 609 a drastic change occurs. At this moment the tightly packed cluster of species begins to fan out in all three
 610 state variables. Simultaneously, we observe large a shift in mean traits for higher values and in additive
 611 genetic variances for lower values. Upon inspecting our calculations, we diagnose the reason for this shift. The
 612 outlier species that were initially pushed away from the common abiotic optimum (0.0 in this case) evolved
 613 a significant reduction in the quantity of heritable variation ($\approx 60\%$) due to directional selection induced
 614 by competition. This reduction in heritable variation slowed adaptation, causing these species to linger on
 615 the outskirts of niche space, some longer than others. In the meantime the rest of the community, being
 616 tightly packed, experienced greater competition which led to diminished abundances for these species and
 617 caused some members of the core group to veer away from the abiotic optimum. The reduced abundances of
 618 the core group led to reduced competition overall. As a result, the outlier populations were given a slight
 619 increase in growth rate, enough to allow them to increase their abundances orders of magnitude higher than

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 |

the species in the core group and giving them more weight in driving the evolution of other species. Many of these heavy-hitting outlier species had already been maintaining negative mean traits, but around time 500.0 the high abundance species with positive mean traits began to experience enough intraspecific competition to override interspecific competition and begin evolving towards the abiotic optimum. The sudden imbalance of these high abundance species effectively induced a single large competitive exclusion event pushing the majority of the community far away from the abiotic optimum. After this shift the cluster began to slowly bloom in all three state variables as species took advantage of novel asymmetries in their competitive abilities mediated by a new distribution of mean trait values across the community. About 125.0 units of time later, the community reached a qualitatively new phase of dynamics. If we kept running the numerical integrator, we would continue to see similar drama unfolding over and over again as minute stochastic changes contribute to asymmetries which slowly build into drastic shifts.

The strong competition scenario is not quite as showy. Although the dynamics of trait means and variances tend to be far more stochastic than in the weak competition scenario, the community overall appears to quickly reach some statistical equilibrium and remain there. However, the abundances across all species in the community are very low due to strength of competition being an order of magnitude higher than in the weak competition case. Although most of the species maintain abundances greater than 1000.0, we found one species that dropped to an abundance of about 50.0. If we let the numerical integrator run long enough in this case, we will likely see many of the species go extinct.

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

3.3 The relation between the strength of ecological interactions and coevolution

647

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

$$\alpha_{ij}(t) = \frac{c_i}{r_i(t)} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p_i(x, t) p_j(y, t) \mathcal{O}_{ij}(x, y) dx dy = \frac{c_i U_i U_j}{r_i(t)} \sqrt{\frac{b_{ij}(t)}{2\pi}} \exp\left(-\frac{b_{ij}(t)}{2} (\bar{x}_i(t) - \bar{x}_j(t))^2\right), \quad (46)$$

650 where

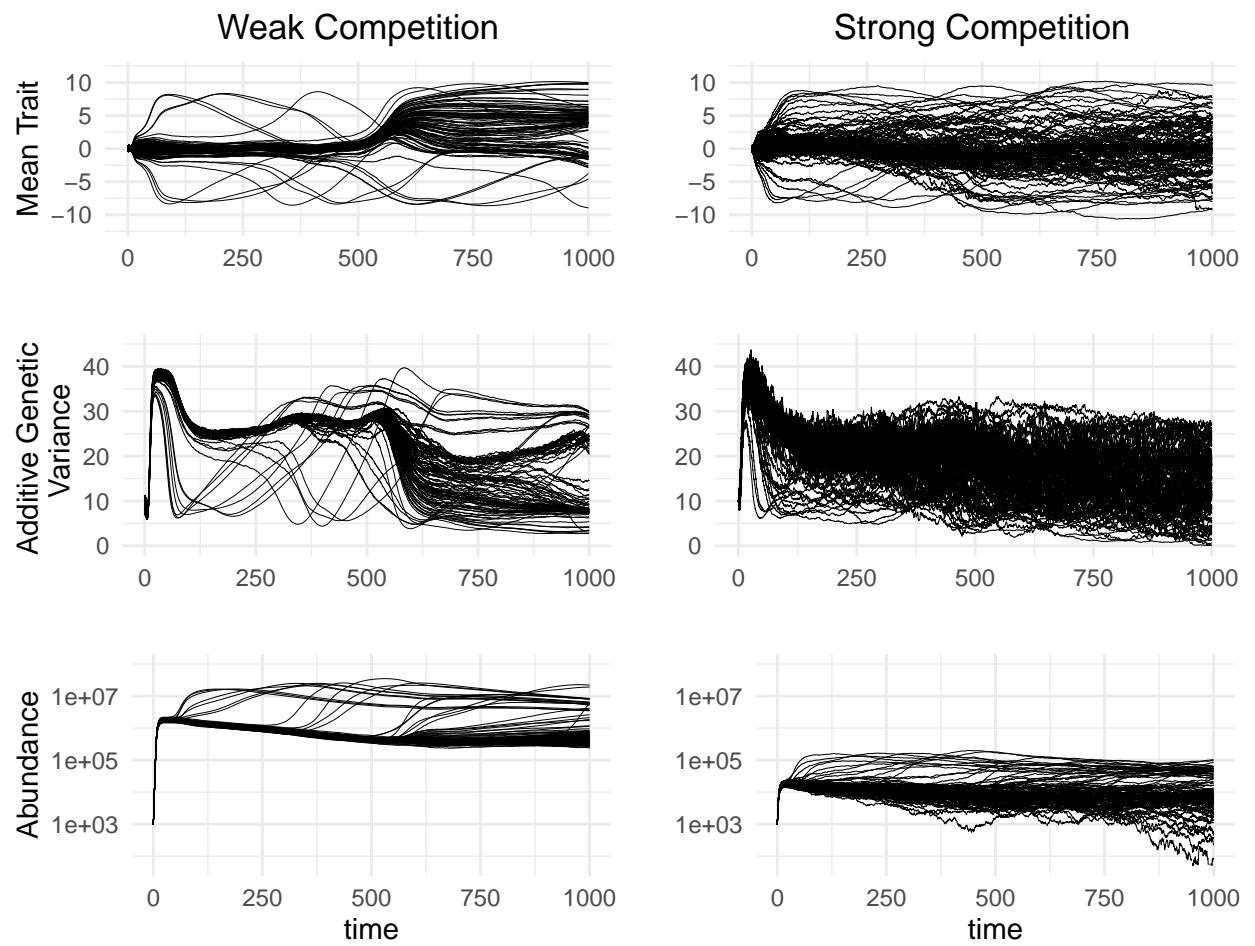


Figure 3: Temporal dynamics of mean trait (top), additive genetic variance (middle) and abundance (bottom) for the scenario of weak competition (left) and strong competition (right). Red lines indicate average trend across species.

$$r_i(t) = R_i - \frac{a_i}{2} \left((\bar{x}_i(t) - \theta_i)^2 + G_i(t) + \eta_i \right). \quad (47)$$

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

$$dN_i(t) = r_i(t) \left(1 - \sum_{j=1}^S \alpha_{ij}(t) N_j(t) \right) N_i(t) dt + \sqrt{V_i N_i(t)} dW_1(t). \quad (48)$$

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

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

$$\beta_{ij}(t) = c_i U_i U_j N_j(t) b_{ij}(t) (\bar{x}_i(t) - \bar{x}_j(t)) \sqrt{\frac{b_{ij}(t)}{2\pi}} \exp \left(-\frac{b_{ij}(t)}{2} (\bar{x}_i(t) - \bar{x}_j(t))^2 \right). \quad (49)$$

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

661 Below we investigate the correspondence of interaction intensity and coevolutionary change. However, we
 662 can already identify one major discrepancy; α_{ij} is maximized when $\bar{x}_i = \bar{x}_j$, but $\beta_{ij} = 0$ under the same
 663 condition. We therefore include in our metric of selection the standardized stabilizing selection gradient γ
 664 which measures the effect of stabilizing or disruptive selection on phenotypic variance (Lande and Arnold
 665 1983). In SM §5.7 we show that the standardized stabilizing selection gradient induced by species j on species
 666 i can be computed as

$$\gamma_{ij}(t) = c_i U_i U_j N_j(t) b_{ij}(t) \left(1 - b_{ij}(t) (\bar{x}_i(t) - \bar{x}_j(t))^2 \right) \sqrt{\frac{b_{ij}(t)}{2\pi}} \exp \left(-\frac{b_{ij}(t)}{2} (\bar{x}_i(t) - \bar{x}_j(t))^2 \right). \quad (50)$$

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

672 We now make use of the expressions derived for competition coefficients and selection gradients to investigate
 673 their relationship. As a first pass, let us assume all model parameters are equivalent across species and that
 674 each species has the same abundance and trait variance. Let us further assume that richness S is large and
 675 the distribution of mean trait values is normal with mean \bar{x} , variance $V_{\bar{x}}$ and density $f_{\bar{x}}$. Such assumptions
 676 are typical when deriving analytical results in the field of theoretical coevolutionary community ecology
 677 (Guimarães, Jordano, and Thompson 2011; Guimarães et al. 2017; Nuismer, Jordano, and Bascompte 2012;
 678 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
 679 approximate r_i with

$$\bar{r} = \int_{-\infty}^{+\infty} \left(R - \frac{a}{2} ((\bar{x} - \theta)^2 + G + \eta) \right) f_{\bar{x}}(\bar{x}) d\bar{x} = R - \frac{a}{2} ((\bar{x} - \theta)^2 + V_{\bar{x}} + G + \eta). \quad (51)$$

680 In SM §5.8 we use these assumptions to calculate the first and second order moments describing the joint
 681 distribution of competition coefficients and selection gradients across the community. We find that the

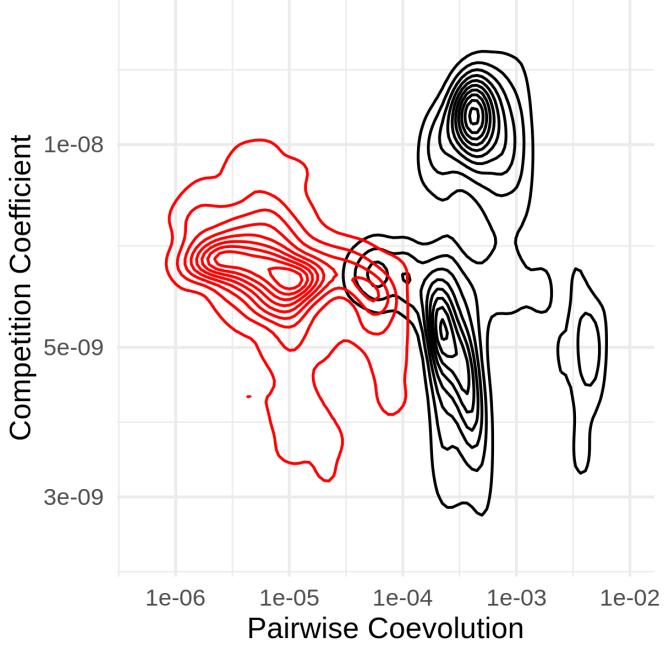


Figure 4: Networks of interspecific interactions parameterized by competition coefficients (top row) and distributions of pairwise coevolution (bottom row) under the scenarios of weak competition (left) and strong competition (right) at time $t = 1.0 \times 10^3$. Node sizes are proportional to population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

682 covariance between linear selection gradients and competition coefficients are zero due to the symmetry
 683 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
 684 covariances between the magnitude of linear selection gradients and competition coefficients and between
 685 stabilizing selection gradients and competition coefficients can be written as

$$\text{Cov}_{f_{\bar{X}}}(\alpha, |\beta|) = \frac{2c^2 b^2 U^4 N}{\pi \bar{r}} \sqrt{\frac{V_{\bar{X}}}{2\pi}} \left(\frac{1}{(1 + 8bV_{\bar{X}})^{3/4}} - \frac{1}{(1 + 4bV_{\bar{X}})^{3/4}(1 + 2bV_{\bar{X}})^{1/2}} \right), \quad (52a)$$

$$\text{Cov}_{f_{\bar{X}}}(\alpha, \gamma) = \frac{c^2 b^2 U^4 N}{2\pi \bar{r}} (1 - 2bV_{\bar{X}}) \left(\frac{1}{\sqrt{1 + 4bV_{\bar{X}}}} - \frac{1}{1 + 2bV_{\bar{X}}} \right). \quad (52b)$$

686 For fixed c, b, \bar{r} and N , we visualize these relationships in Figure 5. To gain insight into the relationship
 687 between selection gradients and competition coefficients, note that our assumptions in this section imply
 688 $b^{-1} = 2(\sigma^2 + w)$. 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
 689 generalists and are comprised of specialist individuals, the value $2bV_{\bar{X}}$ is approximately equal to the ratio of
 690 interspecific mean trait variation to intraspecific individual trait variation. Hence, for both covariances we see
 691 that there is no relationship between selection gradients and competition coefficients when this ratio is zero.
 692 From equation (52a) we can use numerical optimization to find that when $V_{\bar{X}}/\sigma^2 \approx 1.25$ the relationship
 693 between the magnitudes of linear selection gradients and competition coefficients disappears, but when
 694 (approximately) $V_{\bar{X}}/\sigma^2 < 1.25 (> 1.25)$, this covariance becomes negative (positive). Equation (52b) states
 695 that when $V_{\bar{X}}/\sigma^2$ is approximately equal to one (or slightly larger), there is no expected relationship between
 696 competition coefficients and quadratic selection gradients. However, when $V_{\bar{X}}/\sigma^2 < 1.0 (> 1.0)$, then we
 697 expect a positive (negative) relationship between α and γ . These results are true regardless of the chosen
 698 parameter values. In SM §5.8 we use simulations of system (44) to show that these results do not qualitatively
 699 differ when allowing for heterogeneous population sizes and additive genetic variances across species.

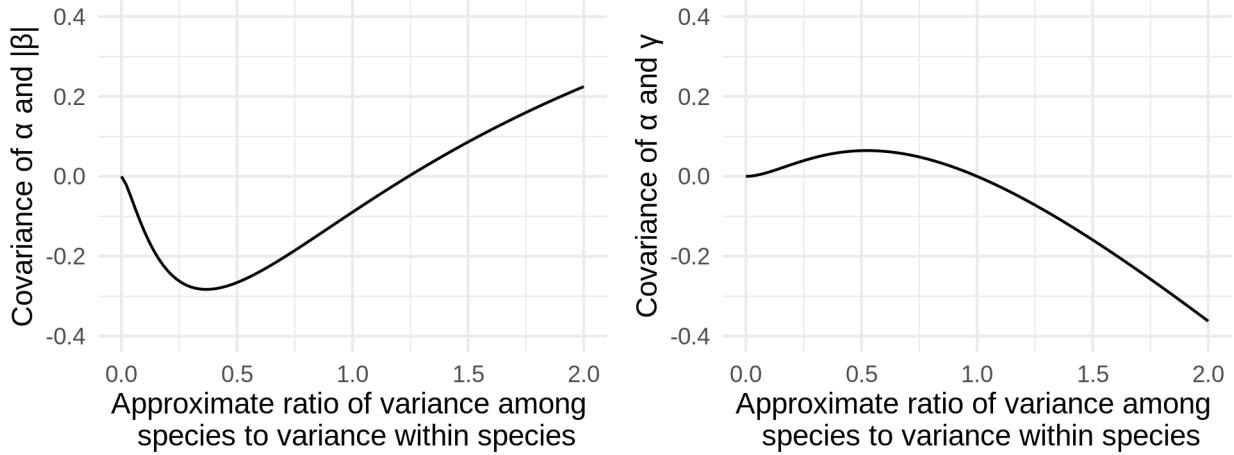


Figure 5: Curves representing the covariance between the magnitude of linear selection gradients and competition coefficients (left) and between stabilizing selection gradients and competition coefficients (right) as a function of $2bV_{\bar{X}}$ which is approximately equal to the ratio of variance 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.

- 700 From a biological perspective, if the ratio $V_{\bar{X}}/\sigma^2$ is small, then species are packed tightly in phenotypic
 701 space. In our model this occurs when abiotic stabilizing selection is much stronger than competition ($a \gg c$).
 702 This causes species to overlap more in niche space (i.e., large α) and creates disruptive selection for greater
 703 intraspecific variance (i.e., positive γ), which explains the positive region of covariance between α and γ .
 704 However, as species begin to overlap in niche space, directional selection begins to vanish (i.e., small $|\beta|$),
 705 leading to a negative covariance between α and $|\beta|$. In the limiting case that two species have perfectly
 706 overlapping niches, they will exhibit zero directional selection since a shift in either direction will yield the
 707 same fitness advantages.
- 708 In the opposite scenario where competition is much stronger than abiotic stabilizing selection ($c \gg a$), species
 709 will not evolve to be as tightly packed and instead their niche-centers will be spread out with little overlap in
 710 their resource utilization curves (i.e., small α). In this case biotic directional selection will be strong (i.e., large
 711 $|\beta|$), particularly for species towards the outer regions of niche space due to asymmetric fitness advantages
 712 conferred by shifts in niche-centers. This leads to a positive covariance between α and $|\beta|$. However, as noted
 713 above, this directional selection will also erode away at standing heritable variation (i.e., negative γ), reducing
 714 the rate at which adaptation can occur and creating a negative covariance between α and γ .
- 715 In summary, we see the relation between competition coefficients and selection is highly non-trivial and
 716 depends on the relative magnitudes of different ecological processes shaping the community. However, this
 717 does not address the relation between competition coefficients and coevolution per se. In SM §5.8 we show
 718 that calculating a formula for the covariance between competition coefficients and the metric of coevolution
 719 \mathfrak{C} introduced above provides a difficult analytical challenge. Instead of confronting this challenge we build
 720 on our numerical approach used to justify analytical approximations of $\text{Cov}_{f_{\bar{X}}}(\alpha, |\beta|)$ and $\text{Cov}_{f_{\bar{X}}}(\alpha, \gamma)$ to
 721 approximate the correlation of α and \mathfrak{C} . This numerical approach inherits the assumptions of homogeneous
 722 background parameters such as the mutation rate μ and abiotic optima θ , but allows us to relax the assumption
 723 that N and G are constant across species and time.
- 724 In particular, we numerically integrated system (44) for $T_1 = 1000.0$ units of time and then continued to
 725 integrate for $T_2 = 1000.0$ units of time. We then calculate the covariance between the quantities α and \mathfrak{C}
 726 across $S = 100$ species for each of the last T_2 time steps. We assume the temporal average of these covariances
 727 across the last T_2 units of time approximates the expectation at equilibrium. We repeated this approach for
 728 randomly drawn a and c until our sample size reached 1000. In Figure 6 we plot the temporally averaged
 729 values of $\text{Cov}_{f_{\bar{X}}}(\alpha, \mathfrak{C})$ against the strength of competition c . Using a cubic regression, we see the correlation

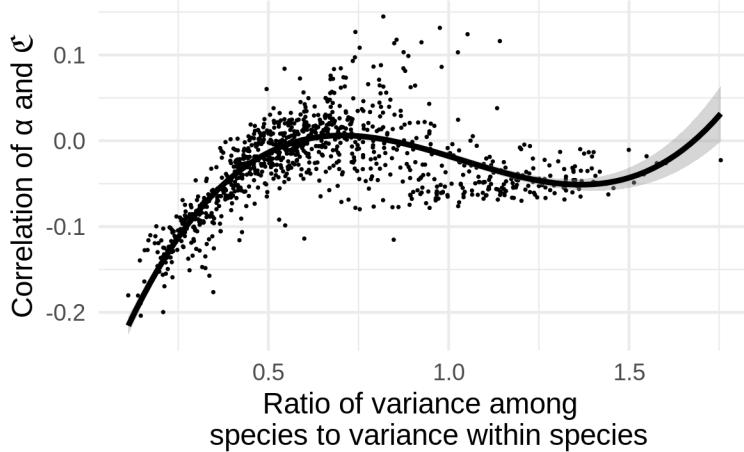


Figure 6: Numerical estimates for the correlation of the strength of coevolution measured by \mathfrak{C} and competition coefficients α plotted against the variance of mean traits among species divided by the mean variance of traits within species. Each dot represents the result from a single simulation.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

730 of coevolutionary selection gradients and competition coefficients is negative at variance ratios below 0.5,
 731 zero at variance ratios between 0.5 and 1.0, and drops below zero again at variance ratios above 1.0.

732 4 Conclusion

733 We have introduced an approach to derive models of evolutionary ecology using the calculus of white noise,
 734 demonstrated our approach by deriving a model of diffuse coevolution and investigated the relationship
 735 between selection gradients and competition coefficients, finding that these quantities exhibit interesting
 736 relationships which shed light on the feedback between the structure and dynamics of ecological communities.
 737 However, although our approach has the merit of rigorously synthesizing the dynamics of abundance and
 738 phenotypic evolution for populations experiencing demographic stochasticity, there remains biological details
 739 and their associated technical challenges that have yet to be confronted for gaining a more thorough and
 740 rigorous understanding of ecological communities. We touch on just four of them here and provide suggestions
 741 for future research to approach these challenges.

742 Limitations of diffusion limits

743 As noted by Feller (1951), although diffusion limits provide reasonable approximations for large populations,
 744 relatively small populations require discrete models. Hence, as a diffusion limit, SPDE (24) restricts parameter
 745 values to regions that maintain large population sizes. Although this suggests an important restriction on
 746 any model developed under this framework by implying populations are not at risk of extinction, the SDE
 747 describing abundance dynamics technically permits extinction. However, for small abundances, pathology
 748 emerges in the evolution of trait means and variances. In particular, stochastic components of the SDE
 749 describing the evolution of \bar{x} and σ^2 diverge towards infinite values as $N \rightarrow 0$. Hence, studies of evolutionary
 750 rescue and colonization-extinction dynamics that incorporate phenotypic evolution should be pursued via a
 751 different approach. A natural alternative can be developed utilizing the underlying BBM that SPDE (24) is a
 752 diffusion limit of (see section 2.2.2). This process explicitly tracks individuals throughout their life-history and
 753 captures reproduction as branching events. Hence, BBM processes model population size as a non-negative
 754 integer instead of a continuously varying number. In particular, the pathological behavior described for the
 755 diffusion limit does not occur for BBM as population size goes to zero.

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

757 Our treatment of inheritance and our approach to model coevolution rest on the assumptions of normally

758 distributed breeding values and expressed phenotypes along with asexual reproduction. However, real traits
759 are not encoded by an infinite number of loci each contributing an additive infinitesimal effect (as required
760 by the infinitesimal model), mutations are not inherited as normally distributed deviations from parental
761 breeding values (as required by the Gaussian descendants model) and many populations of interest exhibit
762 non-random sexual reproduction. Departures from this model of genetic architecture can produce more
763 complex distributions of breeding values and expressed traits. Such deviations can be reinforced via strong
764 non-Gaussian selection surfaces, including the surface $m(\nu, x)$ we have derived from niche theory, along with
765 non-random mating in sexually reproducing populations. However, Gaussian approximations are convenient
766 since they are defined by their mean and variance. Future work investigating the effects of non-normally
767 distributed traits on the structure and dynamics of ecological communities will therefore need to confront
768 higher moments such as skew and kurtosis, ideally integrating previously established mathematical approaches
769 to derive the dynamics of these higher moments (Débarre, Yeaman, and Guillaume 2015).

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

780 **The role of ecological stoichiometry**

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

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

794 **Accounting for macroevolutionary history**

795 To understand patterns found in ecological communities, considerations must push beyond microevolutionary
796 and contemporary ecological processes and consider the macroevolutionary dynamics of ancestral lineages
797 leading to extant populations. Using sub-alpine flower communities as an example, one can observe a very
798 strict ordering of phenology across broad geographic ranges. In particular, whether in the Colorado Rocky
799 mountains (such as Gothic, Colorado) or on an outlier of the Idaho batholith (such as Kamiak butte near
800 Palouse, Washington), one will almost surely observe a very conspicuous order of flowers emerging in early
801 spring: at the very beginning of the season blooms *Claytonia lanceolata* followed by *Erythronium grandiflorum*
802 which precedes *Delphinium nuttallianum* (B. Week, personal observations). If contemporary phenological
803 coevolution is rampant, why should this pattern be so well preserved across a thousand miles of rugged and
804 diverse terrain? Although it would be exciting to find that these species repeatedly coevolved this pattern
805 in each location, a more parsimonious hypothesis suggests the phenology and physiology of these species
806 slowly evolved independently over macroevolutionary time scales to take advantage of the specific conditions
807 available within each of these windows of the flowering season. However, this could not have carried out
808 in the Rocky mountains since this terrain only became habitable just over ten thousand years ago as the

809 glaciers of the Pleistocene began to retreat (Paul CaraDonna, personal communications). Hence, given these
810 considerations, it appears that an understanding of early season phenology patterns should focus on how
811 these communities are assembled as opposed to contemporary evolutionary dynamics. Indeed, recent work
812 testing models of phylogeography ignores the potential for contemporary evolution and instead suggests
813 alpine flower communities tend to follow neutral assembly where flowers merely compete for who can disperse
814 to new habitat first, as opposed to a selective process where a regional species pool is filtered for those species
815 adapted to the newly available habitat (Marx et al. 2017).

816 Of course microevolutionary and ecological dynamics are not completely irrelevant for understanding patterns
817 in communities that are primarily structured by deep evolutionary processes. In particular, macroevolutionary
818 trait evolution is simply the aggregation of microevolutionary change occurring over large spans of time. This
819 suggests a road forward to connect the theory we have introduced to models of macroevolutionary trait
820 evolution.

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

$$dX_t = a(\theta_t - X_t)dt + b dW_t \quad (53)$$

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

843 **Final remarks**

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

856 5 Supplementary material (SM)

857 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
 858 $\Delta = \frac{\partial^2}{\partial x^2}$, except in §5.8.1.3 where Δ represents a random variable.

859 5.1 Sufficient conditions for finite mean, variance and total abundance in the 860 deterministic case

861 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
 862 $\int_{\mathbb{R}} \nu(x, t) dx < +\infty$. 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 $m(\nu, x) \leq r$
 863 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 $\nu(x, 0)$
 864 satisfies

$$N(0) = \int_{\mathbb{R}} \nu(x, 0) dx < +\infty, \quad (54)$$

$$-\infty < \bar{x}(0) = \int_{\mathbb{R}} xp(x, 0) dx < +\infty, \quad (55)$$

$$\sigma^2(0) = \int_{\mathbb{R}} xp(x, 0)^2 dx < +\infty, \quad (56)$$

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

$$\dot{\nu}(x, t) = m(\nu, x)\nu(x, t) + \frac{\mu}{2}\Delta\nu(x, t). \quad (57)$$

868 Replacing m with it's upper bound $r \in \mathbb{R}$, equation (57) reduces to a simple parabolic equation that can be
 869 solved using standard techniques (Farlow 1993). In particular, when $m(\nu, x) \equiv 0$ denote the solution to (57)
 870 by $\nu_0(x, t)$. Then, denoting

$$\Phi(x, t) = \frac{\exp(-x^2/2\mu t)}{\sqrt{2\pi\mu t}}, \quad (58)$$

871 we have

$$\nu_0(x, t) = \int_{\mathbb{R}} \Phi(x - y, t)\nu(y, 0) dy. \quad (59)$$

872 In the more general case, when $m(\nu, x) \equiv r \in \mathbb{R}$, equation (57) has the solution $\nu_r(x, t) = e^{rt}\nu_0(x, t)$.
 873 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
 874 $N_r(t) = \int_{\mathbb{R}} \nu_r(x, t) dx$, $p_r(x, t) = \nu_r(x, t)/N_r(t)$, $\bar{x}_r(t) = \int_{\mathbb{R}} xp_r(x, t) dx$ and $\sigma_r^2(t) = \int_{\mathbb{R}} (x - \bar{x}_r(t))^2 p_r(x, t) dx$,
 875 we have

$$\bar{x}_r(t) = \int_{\mathbb{R}} x \int_{\mathbb{R}} \Phi(x - y, t)p(y, 0) dy dx = \int_{\mathbb{R}} yp(y, 0) dy = \bar{x}(0), \quad (60)$$

$$\sigma_r^2(t) = \int_{\mathbb{R}} (x - \bar{x}_r(t))^2 \int_{\mathbb{R}} \Phi(x - y, t)p(y, 0) dy dx = \int_{\mathbb{R}} ((y - \bar{x}(0))^2 + \mu t) p(y, 0) dy = \sigma^2(0) + \mu t. \quad (61)$$

877 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 $t \geq 0$
 878 such that $\nu(x, t) > \nu_r(x, t)$. Then

$$\nu(x, t) - \nu(x, 0) = \int_0^t m(\nu, x)\nu(x, s) + \frac{\mu}{2}\Delta\nu(x, s) ds > \int_0^t r\nu_r(x, s) + \frac{\mu}{2}\Delta\nu_r(x, s) ds = \nu_r(x, t) - \nu(x, 0) \quad (62)$$

879 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
 880 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 $x \in \mathbb{R}$
 881 and $t \geq 0$. This implies that $N(t) = \int_{\mathbb{R}} \nu(x, t) dx < +\infty$,

$$0 < \int_{\mathbb{R}} x^2 \nu(x, t) dx \leq \int_{\mathbb{R}} x^2 \nu_r(x, t) dx < +\infty \quad (63)$$

882 and in particular

$$0 < \sigma^2(t) + \bar{x}^2(t) = \frac{1}{N(t)} \int_{\mathbb{R}} x^2 \nu(x, t) dx < +\infty \quad (64)$$

883 for each $t \geq 0$.

884 5.2 The relation between diffusion and convolution with a Gaussian kernel

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

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

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

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

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

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

888 Hence, by the fundamental theorem of calculus,

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

889 In particular,

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

890 5.3 Deterministic dynamics of $\sigma^2(t)$

891 Picking up from the main text §2.1,

$$\begin{aligned} \dot{\sigma}^2(t) &= \frac{d}{dt} \int_{\mathbb{R}} (x - \bar{x}(t))^2 p(x, t) dx \\ &= \int_{\mathbb{R}} 2(x - \bar{x}(t)) \dot{x}(t) + (x - \bar{x}(t))^2 \dot{p}(x, t) dx \\ &= \int_{\mathbb{R}} (x - \bar{x}(t))^2 \left((m(\nu, x) - \bar{m}(t)) p(x, t) + \frac{\mu}{2} \frac{\partial^2}{\partial x^2} p(x, t) \right) dx \end{aligned} \quad (65)$$

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

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

$$\frac{\partial^2}{\partial x^2} p(x, t) = \frac{(x - \bar{x}(t))^2 - \sigma^2(t)}{\sigma^4(t)} p(x, t) \quad (66)$$

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

$$\begin{aligned}
&\int_{\mathbb{R}} (x - \bar{x}(t))^2 \frac{\partial^2}{\partial x^2} p(x, t) dx \\
&= \int_{\mathbb{R}} \frac{(x - \bar{x}(t))^4 - (x - \bar{x}(t))^2 \sigma^2(t)}{\sigma^4(t)} p(x, t) dx = 2.
\end{aligned} \quad (67)$$

894 In the Gaussian case, we also have

$$\begin{aligned}
2\sigma^4 \left(\frac{\partial \bar{m}}{\partial \sigma^2} - \overline{\frac{\partial m}{\partial \sigma^2}} \right) &= 2\sigma^4 \left(\frac{\partial}{\partial \sigma^2} \int_{\mathbb{R}} m(\nu, x) p(x, t) dx - \int_{\mathbb{R}} p(x, t) \frac{\partial}{\partial \sigma^2} m(\nu, x) dx \right) \quad (68) \\
&= 2\sigma^4 \int_{\mathbb{R}} \frac{(x - \bar{x})^2 - \sigma^2}{2\sigma^4} m(\nu, x) p(x, t) dx = \int_{\mathbb{R}} ((x - \bar{x})^2 - \sigma^2) (m(\nu, x) - \bar{m}) p(x, t) dx \\
&= \text{Cov}_t((x - \bar{x})^2, m).
\end{aligned}$$

895 5.4 Numerical evidence of finite moments and approximate normality in the 896 stochastic case

897 Here we use a numerical argument to suggest, for

$$r - \frac{a}{2}(\theta - x)^2 - c \int_{\mathbb{R}} \nu(x, t) dx \leq m(\nu, x) \leq r - \frac{a}{2}(\theta - x)^2, \quad \forall (\nu, x) \in C_{1,c}^+(\mathbb{R}) \times \mathbb{R}, \quad (69)$$

898 the density process $\nu(x, t)$ defined by SPDE (24) of the main text satisfies $\int_{\mathbb{R}} (|x| + x^2) \nu(x, t) dx < \infty$. From
899 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

$$\dot{\bar{x}} = aG(\theta - \bar{x}) \quad (70a)$$

$$\dot{G} = \mu - aG^2 \quad (70b)$$

$$\dot{N} = \left(r - \frac{a}{2} ((\theta - \bar{x})^2 + G + \eta) - cN \right) N. \quad (70c)$$

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

$$\hat{\dot{x}} = \theta, \quad (71a)$$

903

$$\hat{G} = \sqrt{\frac{\mu}{a}}, \quad (71b)$$

904

$$\hat{N} = \frac{1}{c} \left(r - \frac{1}{2}(\eta a + \sqrt{\mu a}) \right), \quad (71c)$$

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

911 5.4.1 Description of simulation

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

$$\nu(x, t) = \sum_{i=1}^{N(t)} \delta(x - \xi_i(t)), \quad (72)$$

919 where $\delta(x - \xi_i(t))$ denotes the point mass located at $\xi_i(t)$. To allow for imperfect heritability, we also track
920 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
921 with the quadratic selection gradients discussed in §?? of the main text. Then the distribution of breeding
922 values can be written as

$$\rho(g, t) = \sum_{i=1}^{N(t)} \delta(g - \gamma_i(t)). \quad (73)$$

923 Following our model of heritability, the trait value $\xi_i(t)$ is drawn from a normal distribution centered on
924 $\gamma_i(t)$ with variance η . At each iteration we draw, for each individual, a random number of offspring from
925 a Negative-Binomial distribution. Recall the Negative-Binomial distribution models the number of failed
926 Bernoulli trials that occur before a given number of successful trials. Denoting q the probability of success
927 for each trial and s the number of successes, the mean and variance is given respectively by

$$\frac{s(1-q)}{q}, \quad \frac{s(1-q)}{q^2}. \quad (74)$$

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

$$q(\nu, \xi_i) = \frac{\mathcal{W}(\nu, \xi_i)}{V}, \quad s(\nu, \xi_i) = \frac{\mathcal{W}^2(\nu, \xi_i)}{V - \mathcal{W}(\nu, \xi_i)}. \quad (75)$$

930 The imposes the restriction $V > \mathcal{W}(\nu, \xi_i)$. For each offspring produced by the individual with breeding value
931 $\gamma_i(t)$, 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

$$\mathcal{W}(\nu, \xi_i) = \exp \left(r - \frac{a}{2} (\theta - \xi_i)^2 - c \int_{\mathbb{R}} \nu(x, t) dx \right), \quad (76)$$

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

$$\mathcal{W}(\nu, \xi_i) \rightarrow \mathcal{W}^{(n)}(\nu, \xi_i) = \exp \left(\frac{r}{n} - \frac{a}{2n} (\theta - \xi_i)^2 - \frac{c}{n^2} N(t) \right) = \exp \left(\frac{r}{n} - \frac{a}{2n} (\theta - \xi_i)^2 - \frac{c}{n} N^{(n)}(t) \right). \quad (77)$$

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

$$\hat{N} = \frac{n^2}{c} \left(\frac{r}{n} - \frac{1}{2n} (\eta a + \sqrt{\mu a}) \right) = \frac{n}{c} \left(r - \frac{1}{2} (\eta a + \sqrt{\mu a}) \right). \quad (78)$$

The deterministic equilibrium of the rescaled abundance is then

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

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

$$\lim_{n \rightarrow \infty} n \left(\mathcal{W}^{(n)}(\nu, \xi_i) - 1 \right) = r - \frac{a}{2} (\theta - \xi_i)^2 - c N^{(\infty)}(t). \quad (80)$$

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 $D_n = \sup_x |F_n(x) - F(x)|$.

958 **5.5 Derivation of SDE for \bar{x} and σ^2**

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

$$960 \quad \mathbb{E} \left(\int_{\mathbb{R}} |h(x)| \nu(x, t) dx \right) < \infty, \quad (81)$$

$$961 \quad \mathbb{E} \left(\int_{\mathbb{R}} h^2(x) \nu(x, t) dx \right) < \infty, \quad (82)$$

$$962 \quad \mathbb{E} \left(\int_0^t \int_{\mathbb{R}} \nu(x, s) \left| h(x) \right| \left| m(\nu, x) f(x) + \frac{\mu}{2} \Delta f(x) \right| dx ds \right) < \infty, \quad \forall f \in C_b^2(\mathbb{R}). \quad (83)$$

962 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
963 arguments and integrable in x ,

$$964 \quad \star := \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E}[H(t + \varepsilon) - H(t) | \nu(x, t) = \nu_1(x, t)] \quad (84)$$

$$965 \quad = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E} \left[\int_{\mathbb{R}} h(x) (\nu(x, t + \varepsilon) - \nu(x, t)) dx | \nu(x, t) = \nu_1(x, t) \right] \quad (85)$$

$$966 \quad = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E} \left[\int_{\mathbb{R}} \int_t^{t+\varepsilon} f(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) + \sqrt{V \nu(x, s)} \dot{W}(x, s) \right) ds dx | \nu(x, t) = \nu_1(x, t) \right]. \quad (86)$$

966 By assumption (83) we can use Fubini's theorem to write, with probability one,

$$967 \quad \int_{\mathbb{R}} \int_t^{t+\varepsilon} h(x) (m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s)) ds dx = \int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) (m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s)) dx ds. \quad (87)$$

967 By assumption (82) 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
968 true by definition;

$$969 \quad \int_{\mathbb{R}} \int_t^{t+\varepsilon} h(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) ds dx = \int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) dx ds. \quad (88)$$

969 Hence,

$$970 \quad \mathbb{E} \left[\int_{\mathbb{R}} \int_t^{t+\varepsilon} h(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) ds dx | \nu(x, t) = \nu_1(x, t) \right] \\ = \mathbb{E} \left[\int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) dx ds | \nu(x, t) = \nu_1(x, t) \right] = 0 \quad (89)$$

970 and

$$971 \quad \star = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E} \left[\int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) \right) dx ds | \nu(x, t) = \nu_1(x, t) \right]. \quad (90)$$

971 By assumption (83) we know that there exists a $\delta > 0$ such that for each positive $\varepsilon < \delta$ the following holds
972 almost surely:

$$\begin{aligned} & \left| \int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) \right) dx ds \right| \\ & \leq \int_t^{t+\delta} \int_{\mathbb{R}} |h(x)| \left| m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) \right| dx ds < \infty. \end{aligned} \quad (91)$$

973 Thus, by Lebesgue's dominated convergence theorem, the drift component of the process $H(t)$ can be
974 computed as

$$\begin{aligned} \star = \mathbb{E} \left[\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) \right) dx ds \middle| \nu(x, t) = \nu_1(x, t) \right] \\ = \int_{\mathbb{R}} h(x) \left(m(\nu_1, x) \nu_1(x, t) + \frac{\mu}{2} \Delta \nu_1(x, t) \right) dx. \end{aligned} \quad (92)$$

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

$$\begin{aligned} \star \star := \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{V}[H(t + \varepsilon) - H(t) | \nu(x, t) = \nu_1(x, t)] \\ = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{V} \left[\int_{\mathbb{R}} h(x) (\nu(x, t + \varepsilon) - \nu(x, t)) dx \middle| \nu(x, t) = \nu_1(x, t) \right]. \end{aligned} \quad (93)$$

976 We can rewrite the integral inside expression (93) as

$$\int_{\mathbb{R}} \int_t^{t+\varepsilon} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) + \sqrt{V \nu(x, s)} \dot{W}(x, s) \right) ds dx. \quad (94)$$

977 We have already found

$$\begin{aligned} & \mathbb{E} \left[\int_{\mathbb{R}} \int_t^{t+\varepsilon} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) + \sqrt{V \nu(x, s)} \dot{W}(x, s) \right) ds dx \middle| \nu(x, t) = \nu_1(x, t) \right] \\ & = \mathbb{E} \left[\int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) \right) dx ds \middle| \nu(x, t) = \nu_1(x, t) \right]. \end{aligned} \quad (95)$$

978 Then, since $h^2(x) \nu(x, t)$ is integrable by assumption (82), we have

$$\begin{aligned} \star \star = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E} \left[\left(\int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) ds dx \right)^2 \middle| \nu(x, t) = \nu_1(x, t) \right] \\ = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E} \left[\int_t^{t+\varepsilon} \int_{\mathbb{R}} V h^2(x) \nu(x, s) dx ds \middle| \nu(x, t) = \nu_1(x, t) \right]. \end{aligned} \quad (96)$$

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

$$\int_t^{t+\varepsilon} \int_{\mathbb{R}} V h^2(x) \nu(x, s) dx ds \leq \int_t^{t+\delta} \int_{\mathbb{R}} V h^2(x) \nu(x, s) dx ds. \quad (97)$$

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

$$\star\star = \mathbb{E} \left[\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_t^{t+\varepsilon} \int_{\mathbb{R}} V h^2(x) \nu(x, s) dx ds \middle| \nu(x, t) = \nu_1(x, t) \right] = \int_{\mathbb{R}} V h^2(x) \nu_1(x, t) dx. \quad (98)$$

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

$$dH(t) = \left(\int_{\mathbb{R}} h(x) \left(m(\nu, x) \nu(x, t) + \frac{\mu}{2} \Delta \nu(x, t) \right) dx \right) dt + \sqrt{V \int_{\mathbb{R}} h^2(x) \nu(x, t) dx} dW(t) \quad (99)$$

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

984 5.5.1 Derivation for trait mean

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

$$\begin{aligned} \|N\|_2 &= \sqrt{V \int_{\mathbb{R}} \nu(x, t) dx} = \sqrt{VN} \\ \|\tilde{x}\|_2 &= \sqrt{V \int_{\mathbb{R}} x^2 \nu(x, t) dx} \\ \langle \tilde{x}, N \rangle &= V \int_{\mathbb{R}} x \nu(x, t) dx = \bar{x} VN. \end{aligned} \quad (100)$$

986 Applying formula (99) provides

$$d\tilde{x} = \left(\bar{x}mN + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta \nu(x, t) dx \right) dt + \|\tilde{x}\|_2 d\tilde{W}_2, \quad (101)$$

987 where,

$$d\tilde{W}_2 = d\hat{\mathbf{W}}_{\sqrt{Vx^2\nu}} = \frac{1}{\|\tilde{x}\|_2} \int_{\mathbb{R}} x \sqrt{V\nu(x, t)} \dot{W}(x, t) dx dt. \quad (102)$$

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

$$d\bar{x} = d\left(\frac{\tilde{x}}{N}\right) = \frac{\tilde{x}}{N} \left(\frac{d\tilde{x}}{\tilde{x}} - \frac{dN}{N} - \frac{d\tilde{x}}{\tilde{x}} \frac{dN}{N} + \left(\frac{dN}{N}\right)^2 \right) = \frac{d\tilde{x}}{N} - \bar{x} \frac{dN}{N} - \frac{d\tilde{x}}{N} \frac{dN}{N} + \bar{x} \left(\frac{dN}{N}\right)^2. \quad (103)$$

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

$$\begin{aligned} d\bar{x} &= \bar{x}m dt + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x, t) dx dt + \frac{\|\tilde{x}\|_2}{N} d\tilde{W}_2 - \bar{x} \left(\bar{m} dt + \sqrt{\frac{V}{N}} dW_1 \right) - \frac{\langle \tilde{x}, N \rangle}{N^2} dt + \bar{x} \frac{\|N\|_2^2}{N^2} dt \\ &= (\bar{x}m - \bar{x}\bar{m}) dt + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x, t) dx dt + \frac{\|\tilde{x}\|_2}{N} d\tilde{W}_2 - \bar{x} \sqrt{\frac{V}{N}} dW_1 - V \frac{\bar{x}}{N} dt + V \frac{\bar{x}}{N} dt \\ &= \left(\text{Cov}_t(x, m) + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x, t) dx \right) dt + \frac{\|\tilde{x}\|_2}{N} d\tilde{W}_2 - \bar{x} \sqrt{\frac{V}{N}} dW_1. \end{aligned} \quad (104)$$

⁹⁹⁰ Note that

$$\begin{aligned} \frac{\|\tilde{x}\|_2}{N} d\tilde{W}_2 - \bar{x} \sqrt{\frac{V}{N}} dW_1 &= \frac{1}{N} \int_{\mathbb{R}} x \sqrt{V\nu(x,t)} \dot{W}(x,t) dx - \frac{\bar{x}}{N} \int_{\mathbb{R}} \sqrt{V\nu(x,t)} \dot{W}(x,t) dx \\ &= \int_{\mathbb{R}} \frac{x - \bar{x}}{N} \sqrt{V\nu(x,t)} \dot{W}(x,t) dx \end{aligned} \quad (105)$$

⁹⁹¹ and

$$\mathbb{V} \left(\int_{\mathbb{R}} \frac{x - \bar{x}}{N} \sqrt{V\nu(x,t)} \dot{W}(x,t) dx \right) = \frac{V}{N} \int_{\mathbb{R}} (x - \bar{x})^2 p(x,t) dx = V \frac{\sigma^2}{N}. \quad (106)$$

⁹⁹² Hence, by setting

$$dW_2 = \frac{\int_{\mathbb{R}} \frac{(x - \bar{x})}{N} \sqrt{V\nu(x,t)} \dot{W}(x,t) dx}{\sqrt{V\sigma^2/N}} \quad (107)$$

⁹⁹³ we can write

$$d\bar{x} = \left(\text{Cov}_t(x, m) + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x,t) dx \right) dt + \sqrt{V \frac{\sigma^2}{N}} dW_2. \quad (108)$$

⁹⁹⁴ 5.5.2 Derivation for trait variance

⁹⁹⁵ We set $\tilde{\sigma}^2(t) = \int_{\mathbb{R}} x^2 \nu(x,t) dx$ and make use of the notation

$$\begin{aligned} \|\tilde{\sigma}^2\|_2 &= \sqrt{V \int_{\mathbb{R}} x^4 \nu(x,t) dx} \\ \langle \tilde{\sigma}^2, N \rangle &= V \int_{\mathbb{R}} x^2 \nu(x,t) dx = \bar{x}^2 VN. \end{aligned} \quad (109)$$

⁹⁹⁶ Applying formula (99) provides

$$d\tilde{\sigma}^2 = \left(\bar{x}^2 m N + \frac{\mu}{2} \int_{\mathbb{R}} x^2 \Delta \nu(x,t) dx \right) dt + \|\tilde{\sigma}^2\|_2 d\tilde{W}_3 \quad (110)$$

⁹⁹⁷ where

$$d\tilde{W}_3 = d\hat{\mathbf{W}}_{\sqrt{Vx^4\nu}} = \frac{1}{\|\tilde{\sigma}^2\|_2} \int_{\mathbb{R}} x^2 \sqrt{V\nu(x,t)} \dot{W}(x,t) dx. \quad (111)$$

⁹⁹⁸ Using Itô's quotient rule on $\bar{x}^2 = \tilde{\sigma}^2/N$, we obtain

$$\bar{x}^2 = d \left(\frac{\tilde{\sigma}^2}{N} \right) = \frac{\tilde{\sigma}^2}{N} \left(\frac{d\tilde{\sigma}^2}{\tilde{\sigma}^2} - \frac{dN}{N} - \frac{d\tilde{\sigma}^2}{\tilde{\sigma}^2} \frac{dN}{N} + \left(\frac{dN}{N} \right)^2 \right) = \frac{d\tilde{\sigma}^2}{N} - \bar{x}^2 \frac{dN}{N} - \frac{d\tilde{\sigma}^2}{N} \frac{dN}{N} + \bar{x}^2 \left(\frac{dN}{N} \right)^2. \quad (112)$$

⁹⁹⁹ Table 1 of the main text implies $d\tilde{W}_3 dW_1 = \langle \tilde{\sigma}^2, N \rangle$ and hence

$$\begin{aligned}
d\bar{x}^2 &= \left(\bar{x}^2 \bar{m} + \frac{\mu}{2} \int_{\mathbb{R}} x^2 \Delta p(x, t) dx \right) dt + \frac{\|\tilde{\sigma}^2\|_2}{N} d\tilde{W}_3 - \bar{x}^2 \left(\bar{m} dt + \sqrt{\frac{V}{N}} dW_1 \right) - \frac{\langle \tilde{\sigma}^2, N \rangle}{N^2} dt + \bar{x}^2 \frac{\|N\|_2^2}{N^2} dt \\
&= \left(\bar{x}^2 \bar{m} - \bar{x}^2 \bar{m} dt + \frac{\mu}{2} \int_{\mathbb{R}} x^2 \Delta p(x, t) dx \right) dt + \frac{\|\tilde{\sigma}^2\|_2}{N} d\tilde{W}_3 - \bar{x}^2 \sqrt{\frac{V}{N}} dW_1 - \bar{x}^2 \frac{V}{N} dt + \bar{x}^2 \frac{V}{N} dt \\
&= \left(\text{Cov}_t(x^2, m) + \frac{\mu}{2} \int_{\mathbb{R}} x^2 \Delta p(x, t) dx \right) dt + \frac{\|\tilde{\sigma}^2\|_2}{N} d\tilde{W}_3 - \bar{x}^2 \sqrt{\frac{V}{N}} dW_1. \quad (113)
\end{aligned}$$

1000 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:

$$\begin{aligned}
d\sigma^2 &= d\bar{x}^2 - 2\bar{x}d\bar{x} - (d\bar{x})^2 = \left(\text{Cov}_t(x^2, m) + \frac{\mu}{2} \int_{\mathbb{R}} x^2 \Delta p(x, t) dx \right) dt + \frac{\|\tilde{\sigma}^2\|_2}{N} d\tilde{W}_3 - \bar{x}^2 \sqrt{\frac{V}{N}} dW_1 \\
&- 2\bar{x} \left(\text{Cov}_t(x, m) dt + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x, t) dx dt + \sqrt{\frac{V\sigma^2}{N}} dW_2 \right) - \left(\text{Cov}_t(x, m) dt + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x, t) dx dt + \sqrt{\frac{V\sigma^2}{N}} dW_2 \right)^2 \\
&= \left(\text{Cov}_t(x^2 - 2\bar{x}x, m) + \frac{\mu}{2} \int_{\mathbb{R}} (x^2 - x\bar{x}) \Delta p(x, t) dx \right) dt + \frac{\|\tilde{\sigma}^2\|_2}{N} d\tilde{W}_3 - \bar{x}^2 \sqrt{\frac{V}{N}} dW_1 - 2\bar{x} \sqrt{\frac{V\sigma^2}{N}} dW_2 - \left(\frac{V\sigma^2}{N} \right) dt \\
&= \left(\text{Cov}_t((x - \bar{x})^2, m) + \frac{\mu}{2} \int_{\mathbb{R}} (x - \bar{x})^2 \Delta p(x, t) dx - \frac{V\sigma^2}{N} \right) dt + \frac{\|\tilde{\sigma}^2\|_2}{N} d\tilde{W}_3 - \bar{x}^2 \sqrt{\frac{V}{N}} dW_1 - 2\bar{x} \sqrt{\frac{V\sigma^2}{N}} dW_2. \quad (114)
\end{aligned}$$

1001 In light of

$$\begin{aligned}
\frac{\|\tilde{\sigma}^2\|_2}{N} d\tilde{W}_3 - \bar{x}^2 \sqrt{\frac{V}{N}} dW_1 - 2\bar{x} \sqrt{\frac{V\sigma^2}{N}} dW_2 &= \frac{1}{N} \int_{\mathbb{R}} (x^2 - \sigma^2 - 2\bar{x}(x - \bar{x})) \sqrt{V\nu(x, t)} \dot{W}(x, t) dx \\
&= \frac{1}{N} \int_{\mathbb{R}} ((x - \bar{x})^2 - \sigma^2) \sqrt{V\nu(x, t)} \dot{W}(x, t) dx \quad (115)
\end{aligned}$$

1002 and

$$\begin{aligned}
\frac{1}{N} \int_{\mathbb{R}} \left(((x - \bar{x})^2 - \sigma^2) \sqrt{V\nu(x, s)} \right)^2 dx &= \frac{V}{N} \left(\int_{\mathbb{R}} ((x - \bar{x})^4 - 2(x - \bar{x})^2 \sigma^2 + \sigma^4) p(x, t) dx \right) \\
&= \frac{V}{N} \left(\overline{(x - \bar{x})^4} - \sigma^4 \right) \quad (116)
\end{aligned}$$

1003 we set

$$dW_3 = \frac{\int_{\mathbb{R}} ((x - \bar{x})^2 - \sigma^2) \sqrt{V\nu(x, t)} \dot{W}(x, t) dx}{V \left(\overline{(x - \bar{x})^4} - \sigma^4 \right)} \quad (117)$$

1004 so that

$$d\sigma^2 = \left(\text{Cov}_t((x - \bar{x})^2, m) + \frac{\mu}{2} \int_{\mathbb{R}} (x - \bar{x})^2 \Delta p(x, t) dx - V \frac{\sigma^2}{N} \right) dt + \sqrt{V \frac{(x - \bar{x})^4 - \sigma^4}{N}} dW_3. \quad (118)$$

1005 Table 1 of the main text implies

$$dW_1 dW_2 = \frac{\int_{\mathbb{R}} (x - \bar{x}) \nu(x, t) dx}{\sqrt{N \sigma^2}} dt = 0, \quad (119)$$

$$dW_1 dW_3 = \frac{\int_{\mathbb{R}} ((x - \bar{x})^2 - \sigma^2) \nu(x, t) dx}{\sqrt{(x - \bar{x})^4 - \sigma^4}} dt = 0, \quad (120)$$

$$dW_2 dW_3 = \frac{\int_{\mathbb{R}} (x - \bar{x}) ((x - \bar{x})^2 - \sigma^2) p(x, t) dx}{\sqrt{\sigma^2 ((x - \bar{x})^4 - \sigma^4)}} dt = \frac{N \overline{(x - \bar{x})^3}}{\sqrt{\sigma^2 ((x - \bar{x})^4 - \sigma^4)}} dt. \quad (121)$$

1006 In particular, when p is a Gaussian curve $dW_2 dW_3 = 0$.

1007 5.6 Relating fitness of expressed traits to fitness of breeding values

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

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

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

1008 5.7 Derivation of diffuse coevolution model

1009 In this section we provide a derivation of our model of diffuse coevolution driven by competition. Since most
1010 of the work in this derivation has already been completed in Supplementary Material §5.5, we focus here
1011 on deriving the Malthusian fitness m as a function of trait value x . We begin with discrete populations of
1012 individuals. In particular, we begin by assuming population size n_i is an integer for each species $i = 1, \dots, S$
1013 before passing to the large population size limit.

1014 The reduction in fitness for an individual of species i caused by competition is captured multiplicatively
1015 by $0 < C_i \leq 1$. Biologically this assumes all competitors affect individuals of a given species equally by
1016 consuming the same amount of resources. This is a mean-field interaction since any individual that consumes
1017 resources has an effect on the fitness of all other individuals competing for the same resources. Denote by x_{ij}
1018 the trait value of the j -th individual belonging to species i . The set of trait values across all individuals in the
1019 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
1020 effect of all competitive interactions on the fitness of the j -th individual in species i . Since individuals do
1021 not compete with themselves the net multiplicative effects on fitness of both interspecific and intraspecific
1022 competition on the j -th individual in species i can be summarized by

$$\mathcal{B}_{ij}(X) = C_i^{\sum_{l \neq j} \mathcal{O}_{ii}(x_{ij}, x_{il}) + \sum_{k \neq i} \sum_{l=1}^{n_k} \mathcal{O}_{ik}(x_{ij}, x_{kl})}, \quad (122)$$

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

$$\mathcal{W}_{ij}(X) = \mathcal{A}_i(x_{ij}) \mathcal{B}_{ij}(X), \quad (123)$$

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

$$\mathcal{W}_{1j}^{(\mathbf{n})}(X) = \mathcal{A}_1(x_{1j})^{1/n_1} \exp \left(\frac{\ln C_1}{n_1^2} \sum_{l \neq j} \mathcal{O}_{11}(x_{1j}, x_{1l}) + \frac{\ln C_1}{n_1} \sum_{k \neq 1} \frac{1}{n_k} \sum_{l=1}^{n_k} \mathcal{O}_{1k}(x_{1j}, x_{kl}) \right). \quad (124)$$

For large n_1 , we have the approximation

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

Hence

$$\lim_{n_1 \rightarrow \infty} n_1 (\mathcal{W}_{1j}^{(\mathbf{n})}(X) - 1) = \ln \mathcal{A}_1(x_{1j}) + \left(\int_{\mathbb{R}} \mathcal{O}_{11}(x_{1j}, y) \nu_1(y, t) dy + \sum_{k \neq 1} \frac{1}{n_k} \sum_{l=1}^{n_k} \mathcal{O}_{1k}(x_{1j}, x_{kl}) \right) \ln C_1. \quad (126)$$

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

$$m_1(\boldsymbol{\nu}, x_{1j}) := \lim_{\mathbf{n} \rightarrow \infty} n_1 (\mathcal{W}_{1j}^{(\mathbf{n})}(X) - 1) = \ln \mathcal{A}_1(x) + \left(\sum_{k=1}^S \int_{\mathbb{R}} \mathcal{O}_{1k}(x_{1j}, y) \nu_k(y, t) dy \right) \ln C_1. \quad (127)$$

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

$$\bar{\mathcal{O}}_{ij}(x, t) = \frac{\int_{\mathbb{R}} \mathcal{O}_{ij}(x, y) \nu_j(y, t) dy}{\int_{\mathbb{R}} \nu_j(y, t) dy}. \quad (128)$$

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

$$\bar{\mathcal{O}}_{ij}(x, t) = \frac{\int_{\mathbb{R}} \mathcal{O}_{ij}(x, y) \nu_j(y, t) dy}{\int_{\mathbb{R}} \nu_j(y, t) dy} = \frac{U_i U_j}{\sqrt{2\pi(w_i + w_j + \sigma_j^2(t))}} \exp \left(-\frac{(x - \bar{x}_j(t))^2}{2(w_i + w_j + \sigma_j^2(t))} \right). \quad (129)$$

1039 Setting

$$\sigma_i^2(t) = G_i(t) + \eta_i, \quad (130a)$$

$$R_i = \ln \left(\frac{Q_i U_i}{\sqrt{1 + A_i w_i}} \right), \quad (130b)$$

$$a_i = \frac{A_i}{1 + A_i w_i}, \quad (130c)$$

$$\tilde{b}_{ij}(t) = \frac{1}{w_i + w_j + \sigma_j^2(t)}, \quad (130d)$$

$$c_i = -\ln C_i, \quad (130e)$$

1040 we get

$$m_i(\boldsymbol{\nu}, x) = R_i - \frac{a_i}{2} (x - \theta_i)^2 - c_i \sum_{j=1}^S N_j(t) U_i U_j \sqrt{\frac{\tilde{b}_{ij}(t)}{2\pi}} e^{-\frac{\tilde{b}_{ij}(t)}{2}(x - \bar{x}_j(t))^2}. \quad (131)$$

1041 Hence, our fitness function satisfies condition (??) of the main text.

1042 For the remainder of the derivation we suppress notation indicating dependency on $\boldsymbol{\nu}$, x and t . From (131)
1043 we calculate

$$\frac{\partial m_i}{\partial \bar{x}_i} = c_i N_i U_i^2 \tilde{b}_{ii} (x - \bar{x}_i) \sqrt{\frac{\tilde{b}_{ii}}{2\pi}} e^{-\frac{\tilde{b}_{ii}}{2}(x - \bar{x}_i)^2} \quad (132)$$

$$\begin{aligned} \frac{\partial m_i}{\partial G_i} &= \frac{c_i N_i U_i^2}{2} \left(\frac{(x - \bar{x}_i)^2 - G_i - \eta_i - 2w_i}{(G_i + \eta_i + 2w_i)^2} \right) \sqrt{\frac{\tilde{b}_{ii}}{2\pi}} e^{-\frac{\tilde{b}_{ii}}{2}(x - \bar{x}_i)^2} \\ &= \frac{c_i N_i U_i^2 \tilde{b}_{ii}^2}{2} ((x - \bar{x}_i)^2 - \sigma_i^2 - 2w_i) \sqrt{\frac{\tilde{b}_{ii}}{2\pi}} e^{-\frac{\tilde{b}_{ii}}{2}(x - \bar{x}_i)^2}. \end{aligned} \quad (133)$$

1044 Note that

$$\begin{aligned} &\sqrt{\frac{\tilde{b}_{ii}}{2\pi}} \exp \left(-\frac{\tilde{b}_{ii}}{2}(x - \bar{x}_i)^2 \right) \sqrt{\frac{1}{2\pi\sigma_i^2}} \exp \left(-\frac{(x - \bar{x}_i)^2}{2\sigma_i^2} \right) \\ &= \sqrt{\frac{1}{2\pi(\sigma_i^2 + 1/\tilde{b}_{ii})}} \sqrt{\frac{\sigma_i^2 + 1/\tilde{b}_{ii}}{2\pi\sigma_i^2/\tilde{b}_{ii}}} \exp \left(-\frac{\sigma_i^2 + 1/\tilde{b}_{ii}}{2\sigma_i^2/\tilde{b}_{ii}} (x - \bar{x}_i)^2 \right) \\ &= \sqrt{\frac{1}{4\pi(\sigma_i^2 + w_i)}} \sqrt{\frac{2(\sigma_i^2 + w_i)}{2\pi\sigma_i^2(\sigma_i^2 + 2w_i)}} \exp \left(-\frac{\sigma_i^2(\sigma_i^2 + 2w_i)}{4(\sigma_i^2 + w_i)} (x - \bar{x}_i)^2 \right). \end{aligned} \quad (134)$$

1045 Hence,

$$\overline{\frac{\partial m_i}{\partial \bar{x}_i}} = 0, \quad (135)$$

$$\begin{aligned}\frac{\partial \bar{m}_i}{\partial G_i} &= \frac{c_i N_i U_i^2}{2(\sigma_i^2 + 2w_i)^2} \left(\frac{(\sigma_i^2 + 2w_i)\sigma_i^2}{2(w_i + \sigma_i^2)} - \sigma_i^2 - 2w_i \right) \sqrt{\frac{b_{ii}}{2\pi}} \\ &= \frac{c_i N_i U_i^2}{2(\sigma_i^2 + 2w_i)} \left(\frac{\sigma_i^2}{2(\sigma_i^2 + w_i)} - 1 \right) \sqrt{\frac{b_{ii}}{2\pi}} = -\frac{c_i N_i U_i^2 b_{ii}}{2} \sqrt{\frac{b_{ii}}{2\pi}},\end{aligned}\quad (136)$$

¹⁰⁴⁶ where

$$b_{ij} = \frac{1}{w_i + w_j + \sigma_i^2 + \sigma_j^2}. \quad (137)$$

¹⁰⁴⁷ The average fitness for species i is

$$\bar{m}_i = R_i - \frac{a_i}{2} \left((\bar{x}_i - \theta_i)^2 + G_i + \eta_i \right) - c_i \sum_{j=1}^S N_j U_i U_j \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2}. \quad (138)$$

¹⁰⁴⁸ Thus,

$$\frac{\partial \bar{m}_i}{\partial \bar{x}_i} = a_i(\theta_i - \bar{x}_i) - c_i \sum_j N_j U_i U_j b_{ij} (\bar{x}_j - \bar{x}_i) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2}, \quad (139)$$

$$\frac{\partial \bar{m}_i}{\partial G_i} = -\frac{a_i}{2} + \frac{c_i}{2} \sum_{j=1}^S N_j U_i U_j b_{ij} (1 - b_{ij}(\bar{x}_i - \bar{x}_j)^2) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2}. \quad (140)$$

¹⁰⁴⁹ In particular

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

¹⁰⁵⁰ Applying equations (29a), (39a) and (39b) of the main text recovers system (44) of the main text.

¹⁰⁵¹ 5.8 The relation between competition coefficients and selection

¹⁰⁵² 5.8.1 Derivation of analytical approximations

¹⁰⁵³ Just as with most calculations in this work, the derivations are straightforward applications of Gaussian
¹⁰⁵⁴ products. That is, if

$$f_1(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(\mu_1 - x)^2}{2\sigma_1^2}\right), \quad f_2(x) = \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(\mu_2 - x)^2}{2\sigma_2^2}\right), \quad (142)$$

¹⁰⁵⁵ then

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

¹⁰⁵⁶ where

$$\tilde{\mu} = \frac{\sigma_2^2 \mu_1 + \sigma_1^2 \mu_2}{\sigma_1^2 + \sigma_2^2}, \quad \tilde{\sigma}^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}. \quad (144)$$

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

¹⁰⁵⁸ Recalling

$$\alpha(\bar{x}_i, \bar{x}_j) = \frac{c}{\bar{r}} \sqrt{\frac{b}{2\pi}} \exp\left(-\frac{b}{2} (\bar{x}_i - \bar{x}_j)^2\right), \quad (145)$$

$$\gamma(\bar{x}_i, \bar{x}_j) = cNb \left(1 - b(\bar{x}_i - \bar{x}_j)^2\right) \sqrt{\frac{b}{2\pi}} \exp\left(-\frac{b}{2} (\bar{x}_i - \bar{x}_j)^2\right), \quad (146)$$

¹⁰⁵⁹ we have

$$\begin{aligned} \bar{\alpha} &= \int_{\mathbb{R}} \int_{\mathbb{R}} \alpha(\bar{x}_i, \bar{x}_j) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\ &= \frac{c}{\bar{r}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi(b^{-1} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(b^{-1} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j = \frac{c/\bar{r}}{\sqrt{2\pi(b^{-1} + 2V_{\bar{X}})}}, \end{aligned} \quad (147)$$

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

$$\begin{aligned} \text{Var}_{f_{\bar{X}}}(\alpha) &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{\alpha} - \alpha(\bar{x}_i, \bar{x}_j))^2 f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\ &= \frac{c^2}{\bar{r}^2} \left(\sqrt{\frac{b}{4\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} \sqrt{\frac{b}{\pi}} \exp(-b(\bar{x}_i - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j - \frac{1}{2\pi(b^{-1} + 2V_{\bar{X}})} \right) \\ &= \frac{c^2}{\bar{r}^2} \left(\sqrt{\frac{b}{4\pi}} \int_{\mathbb{R}} \sqrt{\frac{1}{2\pi(\frac{1}{2b} + V_{\bar{X}})}} \exp(-b(\bar{x} - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j - \frac{1}{2\pi(b^{-1} + 2V_{\bar{X}})} \right) \\ &= \frac{c^2}{\bar{r}^2} \left(\sqrt{\frac{b}{4\pi}} \sqrt{\frac{1}{2\pi(\frac{1}{2b} + 2V_{\bar{X}})}} - \frac{1}{2\pi(b^{-1} + 2V_{\bar{X}})} \right) = \frac{c^2 b}{2\pi \bar{r}^2} \left(\frac{1}{\sqrt{1 + 4bV_{\bar{X}}}} - \frac{1}{1 + 2bV_{\bar{X}}} \right), \end{aligned} \quad (149)$$

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

1060 **5.8.1.2 Caclulating $\text{Cov}_{f_{\bar{X}}}(\alpha, |\beta|)$**

1061 To calculate moments of $|\beta|$ we note that, as a random variable, $|\beta|$ takes a folded normal distribution. Setting
1062 $\Phi(x)$ equal to the cumulative density function of the standard normal distribution and using the properties of
1063 the folded normal distribution, we find

$$|\bar{\beta}| = \sqrt{\frac{2\text{Var}_{f_{\bar{X}}}(\beta)}{\pi}} \exp\left(-\frac{\bar{\beta}^2}{2\text{Var}_{f_{\bar{X}}}(\beta)}\right) - \bar{\beta} \left[1 - 2\Phi\left(\frac{\bar{\beta}}{\sqrt{\text{Var}_{f_{\bar{X}}}(\beta)}}\right)\right] \quad (151)$$

$$\text{Var}_{f_{\bar{X}}}(|\beta|) = \bar{\beta}^2 + \text{Var}_{f_{\bar{X}}}(\beta) - |\bar{\beta}|^2. \quad (152)$$

1064 Recall that

$$\beta(\bar{x}_i, \bar{x}_j) = cNb(\bar{x}_i - \bar{x}_j) \sqrt{\frac{b}{2\pi}} \exp\left(-\frac{b}{2}(\bar{x}_i - \bar{x}_j)^2\right) \quad (153)$$

1065 and hence

$$\begin{aligned}
\bar{\beta} &= \int_{\mathbb{R}} \int_{\mathbb{R}} \beta(\bar{x}_i, \bar{x}_j) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= cNb \int_{\mathbb{R}} (\bar{x} - \bar{x}_j) \frac{1}{\sqrt{2\pi(b^{-1} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(b^{-1} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j = 0, \quad (154)
\end{aligned}$$

$$\begin{aligned}
\text{Var}_{f_{\bar{X}}}(\beta) &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{\beta} - \beta(\bar{x}_i, \bar{x}_j))^2 f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} c^2 N^2 b^2 (\bar{x}_i - \bar{x}_j)^2 \frac{b}{2\pi} \exp(-b(\bar{x}_i - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \sqrt{\frac{b}{4\pi}} c^2 N^2 b^2 \int_{\mathbb{R}} \left[\left(\frac{\bar{x} + 2bV_{\bar{X}}\bar{x}_j}{1 + 2bV_{\bar{X}}} - \bar{x}_j \right)^2 + \frac{V_{\bar{X}}}{1 + 2bV_{\bar{X}}} \right] \frac{\exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(\frac{1}{2b} + V_{\bar{X}})}\right)}{\sqrt{2\pi(\frac{1}{2b} + V_{\bar{X}})}} f_{\bar{X}}(\bar{x}_j) d\bar{x}_j \\
&= \sqrt{\frac{b}{4\pi}} c^2 N^2 b^2 \int_{\mathbb{R}} \left[\frac{(\bar{x} - \bar{x}_j)^2}{(1 + 2bV_{\bar{X}})^2} + \frac{V_{\bar{X}}}{1 + 2bV_{\bar{X}}} \right] \frac{\exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(\frac{1}{2b} + V_{\bar{X}})}\right)}{\sqrt{2\pi(\frac{1}{2b} + V_{\bar{X}})}} f_{\bar{X}}(\bar{x}_j) d\bar{x}_j \\
&= \sqrt{\frac{b}{4\pi}} c^2 N^2 b^2 \left[\frac{(1 + 2bV_{\bar{X}})V_{\bar{X}}}{1 + 4bV_{\bar{X}}} \frac{1}{(1 + 2bV_{\bar{X}})^2} + \frac{V_{\bar{X}}}{1 + 2bV_{\bar{X}}} \right] \frac{1}{\sqrt{2\pi(\frac{1}{2b} + 2V_{\bar{X}})}} \\
&= \frac{b}{\pi} \frac{c^2 N^2 b^2}{\sqrt{1 + 4bV_{\bar{X}}}} \frac{V_{\bar{X}}}{1 + 2bV_{\bar{X}}} \left(\frac{1}{1 + 4bV_{\bar{X}}} + 1 \right) = \frac{2c^2 N^2 b^3 V_{\bar{X}}}{\pi(1 + 4bV_{\bar{X}})^{3/2}}. \quad (155)
\end{aligned}$$

1066 Thus, using properties of the folded normal distribution, we find

$$\overline{|\beta|} = \sqrt{\frac{2}{\pi}} \frac{cNb^{3/2}}{(1 + 4bV_{\bar{X}})^{3/4}} \sqrt{\frac{2V_{\bar{X}}}{\pi}} = \frac{2}{\pi} \frac{cNb^{3/2}}{(1 + 4bV_{\bar{X}})^{3/4}} \sqrt{V_{\bar{X}}}, \quad (156)$$

$$\text{Var}_{f_{\bar{X}}}(|\beta|) = \frac{c^2 N^2 b^3}{(1 + 4bV_{\bar{X}})^{3/2}} \frac{2V_{\bar{X}}}{\pi} \left(1 - \frac{2}{\pi} \right). \quad (157)$$

1067 We also calculate

$$\begin{aligned}
\text{Cov}_{f_{\bar{X}}}(\alpha, \beta) &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{\alpha} - \alpha(\bar{x}_i, \bar{x}_j)) (\bar{\beta} - \beta(\bar{x}_i, \bar{x}_j)) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \frac{c^2 Nb}{2\bar{r}} \sqrt{\frac{b}{\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{x}_i - \bar{x}_j) \sqrt{\frac{b}{\pi}} \exp(-b(\bar{x}_i - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j = 0. \quad (158)
\end{aligned}$$

1068 In attempt to calculate $\text{Cov}_{f_{\bar{X}}}(\alpha, |\beta|)$ we find

$$\begin{aligned}
\text{Cov}_{f_{\bar{X}}}(\alpha, |\beta|) &= \int_{\mathbb{R}} \int_{\mathbb{R}} \alpha(\bar{x}_i, \bar{x}_j) |\beta(\bar{x}_i, \bar{x}_j)| f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j - \bar{\alpha} \overline{|\beta|} \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{c}{\bar{r}} \sqrt{\frac{b}{2\pi}} \exp\left(-\frac{b}{2}(\bar{x}_i - \bar{x}_j)^2\right) cNb |\bar{x}_i - \bar{x}_j| \sqrt{\frac{b}{2\pi}} \exp\left(-\frac{b}{2}(\bar{x}_i - \bar{x}_j)^2\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j - \bar{\alpha} \overline{|\beta|} \\
&= \frac{c^2 Nb}{\bar{r}} \sqrt{\frac{b}{4\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} |\bar{x}_i - \bar{x}_j| \sqrt{\frac{b}{\pi}} \exp(-b(\bar{x}_i - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j - \bar{\alpha} \overline{|\beta|}. \quad (159)
\end{aligned}$$

1069 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

$$\int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{x}_i - \bar{x}_j) \sqrt{\frac{b}{\pi}} \exp(-b(\bar{x}_i - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j = 0 \quad (160)$$

¹⁰⁷⁰ and

$$\begin{aligned}
& \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{x}_i - \bar{x}_j)^2 \frac{b}{\pi} \exp(-2b(\bar{x}_i - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \sqrt{\frac{2b}{\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{x}_i - \bar{x}_j)^2 \frac{1}{\sqrt{2\pi^{\frac{1}{4b}}}} \exp\left(-\frac{(\bar{x}_i - \bar{x}_j)^2}{2^{\frac{1}{4b}}}\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \sqrt{\frac{2b}{\pi}} \int_{\mathbb{R}} \left[\left(\frac{\bar{x} + 4bV_{\bar{X}}\bar{x}_j}{1 + 4bV_{\bar{X}}} - \bar{x}_j \right)^2 + \frac{V_{\bar{X}}}{1 + 4bV_{\bar{X}}} \right] \frac{1}{\sqrt{2\pi(\frac{1}{4b} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(\frac{1}{4b} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j \\
&= \sqrt{\frac{2b}{\pi}} \int_{\mathbb{R}} \left[\left(\frac{\bar{x} - \bar{x}_j}{1 + 4bV_{\bar{X}}} \right)^2 + \frac{V_{\bar{X}}}{1 + 4bV_{\bar{X}}} \right] \frac{1}{\sqrt{2\pi(\frac{1}{4b} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(\frac{1}{4b} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j \\
&= \sqrt{\frac{2b}{\pi}} \left[\frac{(1 + 4bV_{\bar{X}})V_{\bar{X}}}{1 + 8bV_{\bar{X}}} \frac{1}{(1 + 4bV_{\bar{X}})^2} + \frac{V_{\bar{X}}}{1 + 4bV_{\bar{X}}} \right] \frac{1}{\sqrt{2\pi(\frac{1}{4b} + 2V_{\bar{X}})}} \\
&= \sqrt{\frac{2b}{\pi}} \frac{2V_{\bar{X}}}{1 + 8bV_{\bar{X}}} \sqrt{\frac{4b}{2\pi(1 + 8bV_{\bar{X}})}} = \frac{b}{\pi} \frac{4V_{\bar{X}}}{(1 + 8bV_{\bar{X}})^{3/2}}. \quad (161)
\end{aligned}$$

¹⁰⁷¹ Hence

$$\int_{\mathbb{R}} \int_{\mathbb{R}} |\bar{x}_i - \bar{x}_j| \sqrt{\frac{b}{\pi}} \exp(-b(\bar{x}_i - \bar{x}_j)^2) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j = \sqrt{\frac{2}{\pi}} \sqrt{\frac{b}{\pi} \frac{4V_{\bar{X}}}{(1 + 8bV_{\bar{X}})^{3/2}}} = \frac{2}{\pi} \frac{\sqrt{2bV_{\bar{X}}}}{(1 + 8bV_{\bar{X}})^{3/2}} \quad (162)$$

¹⁰⁷² and

$$\begin{aligned}
\text{Cov}_{f_{\bar{X}}}(\alpha, |\beta|) &= \frac{c^2 N b}{\bar{r}} \sqrt{\frac{b}{4\pi}} \frac{2}{\pi} \frac{\sqrt{2bV_{\bar{X}}}}{(1 + 8bV_{\bar{X}})^{3/4}} - \bar{\alpha} |\beta| \\
&= \frac{2c^2 N b^2}{\pi \bar{r} (1 + 8bV_{\bar{X}})^{3/4}} \sqrt{\frac{V_{\bar{X}}}{2\pi}} - \frac{c}{\bar{r}} \sqrt{\frac{b}{2\pi(1 + 2bV_{\bar{X}})}} \frac{2}{\pi} \frac{c N b^{3/2}}{(1 + 4bV_{\bar{X}})^{3/4}} \sqrt{V_{\bar{X}}} \\
&= \frac{2c^2 N b^2}{\pi \bar{r} (1 + 8bV_{\bar{X}})^{3/4}} \sqrt{\frac{V_{\bar{X}}}{2\pi}} - \frac{2c^2 N b^2}{\pi \bar{r} (1 + 4bV_{\bar{X}})^{3/4}} \sqrt{\frac{V_{\bar{X}}}{2\pi(1 + 2bV_{\bar{X}})}} \\
&= \frac{2c^2 N b^2}{\pi \bar{r}} \sqrt{\frac{V_{\bar{X}}}{2\pi}} \left(\frac{1}{(1 + 8bV_{\bar{X}})^{3/4}} - \frac{1}{(1 + 4bV_{\bar{X}})^{3/4}(1 + 2bV_{\bar{X}})^{1/2}} \right). \quad (163)
\end{aligned}$$

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

¹⁰⁷⁴ We have

$$\mathfrak{C}(\bar{x}_i, \bar{x}_j) = c^2 N^2 b^2 \left(|\bar{x}_i - \bar{x}_j| + |1 - b(\bar{x}_i - \bar{x}_j)^2| \right)^2 \exp\left(-\frac{b}{2}(\bar{x}_i - \bar{x}_j)^2\right). \quad (164)$$

¹⁰⁷⁵ 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
¹⁰⁷⁶ write the probability density function of δ as $f_{\Delta}(\delta)$. Substituting in δ , we can write

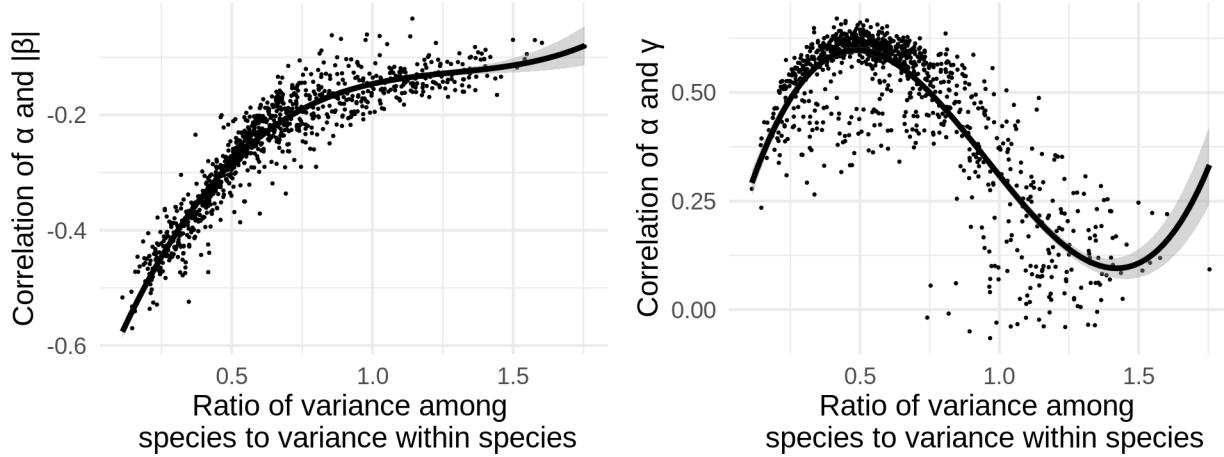


Figure 7: Numerical estimate for the correlations of selection gradients and competition coefficients.

$$\begin{aligned} \mathfrak{C}(\delta, 0) &= c^2 N^2 b^2 \left(|\delta| + |1 - b\delta^2| \right)^2 \exp \left(-\frac{b}{2}\delta^2 \right) \\ &= c^2 N^2 b^2 \left(\delta^2 + 2|\delta| - b|\delta|^3 + (1 - b\delta^2)^2 \right) \exp \left(-\frac{b}{2}\delta^2 \right). \quad (165) \end{aligned}$$

1077 From this expression, we see properties of the folded normal distribution can be used to calculate several
1078 components of the intregal $\text{Cov}_{f_{\bar{X}}}(\alpha, \mathfrak{C})$, but a major technical challenge lies in calculating

$$\int_{\mathbb{R}} | |\delta| - b|\delta|^3 | \exp \left(-\frac{b}{2}\delta^2 \right) f_{\Delta}(\delta) d\delta. \quad (166)$$

1079 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
1080 outlined in the following section.

1081 **5.8.2 Numerical estimates for heterogeneous N and G**

1082 Details on simulations, table of parameters, distributions of a and c .

1083 **References**

- 1084 Abdala-Roberts, Luis, and Kailen A. Mooney. 2014. "Ecological and Evolutionary Consequences of Plant
1085 Genotype Diversity in a Tri-Trophic System." *Ecology* 95 (10). Wiley: 2879–93.
- 1086 Ågren, Göran I., and Folke O. Andersson. 2012. *Terrestrial Ecosystem Ecology: Principles and Applications*.
1087 Cambridge University Press.
- 1088 Barton, N.H., A.M. Etheridge, and A. Véber. 2017. "The Infinitesimal Model: Definition, Derivation, and
1089 Implications." *Theoretical Population Biology* 118 (December). Elsevier BV: 50–73.
- 1090 Barton, Nick, and Alison Etheridge. 2019. "Mathematical Models in Population Genetics." Wiley.
- 1091 Barton, Nick, Alison Etheridge, and Amandine Véber. 2013. "Modelling Evolution in a Spatial Continuum."
1092 *Journal of Statistical Mechanics: Theory and Experiment* 2013 (01). IOP Publishing: P01002.
- 1093 Bertoin, Jean, and Jean-François Le Gall. 2003. "Stochastic Flows Associated to Coalescent Processes."
1094 *Probability Theory and Related Fields* 126 (2). Springer Science; Business Media: 261–88.
- 1095 Bulmer. 1980. *The Mathematical Theory of Quantitative Genetics*. Oxford University Press.
- 1096 Bürger, Reinhard. 1986. "On the Maintenance of Genetic Variation: Global Analysis of Kimuras Continuum-
1097 of-Alleles Model." *Journal of Mathematical Biology* 24 (3). Springer Nature: 341–51.
- 1098 ———. 2000. *The Mathematical Theory of Selection, Recombination, and Mutation*. Wiley.
- 1099 Champagnat, Nicolas, Régis Ferrière, and Sylvie Méléard. 2006. "Unifying Evolutionary Dynamics: From
1100 Individual Stochastic Processes to Macroscopic Models." *Theoretical Population Biology* 69 (3). Elsevier BV:
1101 297–321.
- 1102 Chesson, Peter. 2000. "Mechanisms of Maintenance of Species Diversity." *Annual Review of Ecology and
1103 Systematics* 31 (1). Annual Reviews: 343–66.
- 1104 Conner, Jeffrey K. 2004. *A Primer of Ecological Genetics*. Sinauer Associates.
- 1105 Crow, James F., and Motoo Kimura. 1970. *An Introduction to Population Genetics Theory*. The Blackburn
1106 Press.
- 1107 Crutsinger, Gregory M. 2015. "A Community Genetics Perspective: Opportunities for the Coming Decade."
1108 *New Phytologist* 210 (1). Wiley: 65–70.
- 1109 Da Prato, Giuseppe, and Jerzy Zabczyk. 2014. *Stochastic Equations in Infinite Dimensions*. Cambridge
1110 University Press.
- 1111 David, Thomas I., Jonathan Storkey, and Carly J. Stevens. 2019. "Understanding How Changing Soil
1112 Nitrogen Affects Plant-pollinator Interactions." *Arthropod-Plant Interactions* 13 (5). Springer Science; Business
1113 Media LLC: 671–84.
- 1114 Dawson, Donald A. 1975. "Stochastic Evolution Equations and Related Measure Processes." *Journal of
1115 Multivariate Analysis* 5 (1). Elsevier BV: 1–52.
- 1116 ———. 1993. "Measure-Valued Markov Processes." In *École d'été de Probabilités de Saint-Flour Xxi-1991*,
1117 1–260. Springer.
- 1118 Débarre, Florence, Sam Yeaman, and Frédéric Guillaume. 2015. "Evolution of Quantitative Traits Under a
1119 Migration-Selection Balance: When Does Skew Matter?" *The American Naturalist* 186 Suppl 1: S37–47.
- 1120 Doebeli, Michael. 1996. "Quantitative Genetics and Population Dynamics." *Evolution* 50 (2). Wiley: 532–46.
- 1121 Etheridge, Alison M. 2000. *An Introduction to Superprocesses*. American Mathematical Society.
- 1122 ———. 2008. "Drift, Draft and Structure: Some Mathematical Models of Evolution." In *Stochastic Models in
1123 Biological Sciences*. Institute of Mathematics Polish Academy of Sciences.

- 1124 Etheridge, Alison, and Peter March. 1991. "A Note on Superprocesses." *Probability Theory and Related*
 1125 *Fields* 89 (2). Springer: 141–47.
- 1126 Evans, Lawrence C. 2010. *Partial Differential Equations: Second Edition*. American Mathematical Society.
- 1127 ———. 2014. *An Introduction to Stochastic Differential Equations*. American Mathematical Society.
- 1128 Ewens, Warren J. 2004. *Mathematical Population Genetics*. Springer New York.
- 1129 Falconer, Kenneth. 2014. *Fractal Geometry: Mathematical Foundations and Applications*. Wiley.
- 1130 Farlow, Stanley J. 1993. *Partial Differential Equations for Scientists and Engineers*. Dover.
- 1131 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.
- 1132 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.
- 1133 Fisher, R. A. 1923. "XXI.-on the Dominance Ratio." *Proceedings of the Royal Society of Edinburgh* 42. Cambridge University Press: 321–41.
- 1134 Fitzpatrick, Connor R., Anurag A. Agrawal, Nathan Basiliko, Amy P. Hastings, Marney E. Isaac, Michael Preston, and Marc T. J. Johnson. 2015. "The Importance of Plant Genotype and Contemporary Evolution for Terrestrial Ecosystem Processes." *Ecology* 96 (10). Wiley: 2632–42.
- 1135 Fitzpatrick, Connor R., Anna V. Mikhailichenko, Daniel N. Anstett, and Marc T. J. Johnson. 2017. "The Influence of Range-Wide Plant Genetic Variation on Soil Invertebrate Communities." *Ecography* 41 (7). Wiley: 1135–46.
- 1136 Frank, S. A. 2012. "Natural Selection. IV. The Price Equation." *Journal of Evolutionary Biology* 25 (6). Wiley: 1002–19.
- 1137 Fridley, Jason D. 2017. "Plant Energetics and the Synthesis of Population and Ecosystem Ecology." Edited by David Gibson. *Journal of Ecology* 105 (1). Wiley: 95–110.
- 1138 Fussman, G. F., M. Loreau, and P. A. Abrams. 2007. "Eco-Evolutionary Dynamics of Communities and Ecosystems." *Functional Ecology* 21 (3). Wiley: 465–77.
- 1139 Guimarães, Paulo R., Pedro Jordano, and John N. Thompson. 2011. "Evolution and Coevolution in Mutualistic Networks." *Ecology Letters* 14 (9). Wiley: 877–85.
- 1140 Guimarães, Paulo R., Mathias M. Pires, Pedro Jordano, Jordi Bascompte, and John N. Thompson. 2017. "Indirect Effects Drive Coevolution in Mutualistic Networks." *Nature* 550 (7677). Springer Science; Business Media LLC: 511–14.
- 1141 Harmon, Luke J., Cecilia S. Andreazza, Florence Débarre, Jonathan Drury, Emma E. Goldberg, Ayana B. Martins, Carlos J. Melián, et al. 2019. "Detecting the Macroevolutionary Signal of Species Interactions." *Journal of Evolutionary Biology* 32 (8). Wiley: 769–82.
- 1142 Harte, John. 2011. *Maximum Entropy and Ecology*. Oxford University Press.
- 1143 Harte, John, and Erica A. Newman. 2014. "Maximum Information Entropy: A Foundation for Ecological Theory." *Trends in Ecology & Evolution* 29 (7). Elsevier BV: 384–89.
- 1144 Hickerson, M.J., B.C. Carstens, J. Cavender-Bares, K.A. Crandall, C.H. Graham, J.B. Johnson, L. Rissler, P.F. Victoriano, and A.D. Yoder. 2010. "Phylogeography's Past, Present, and Future: 10 Years After Avise, 2000?" *Molecular Phylogenetics and Evolution* 54 (1). Elsevier BV: 291–301.
- 1145 Hofbauer, Josef, and Karl Sigmund. 1998. *Evolutionary Games and Population Dynamics*. Cambridge University Press.
- 1146 Holt, Robert D. 1987. "On the Relation Between Niche Overlap and Competition: The Effect of Incommensurable Niche Dimensions." *Oikos* 48 (1). JSTOR: 110.

- ¹¹⁶⁷ Kendall, David G. 1966. "Branching Processes Since 1873." *Journal of the London Mathematical Society* s1-41 (1). Wiley: 385–406.
- ¹¹⁶⁹ Kimmel, Marek, and David E. Axelrod. 2015. *Branching Processes in Biology*. Springer New York.
- ¹¹⁷⁰ Kimura, M. 1965. "A Stochastic Model Concerning the Maintenance of Genetic Variability in Quantitative Characters." *Proceedings of the National Academy of Sciences* 54 (3). Proceedings of the National Academy of Sciences: 731–36.
- ¹¹⁷³ Kimura, M., and J. F. Crow. 1978. "Effect of Overall Phenotypic Selection on Genetic Change at Individual Loci." *Proceedings of the National Academy of Sciences* 75 (12). Proceedings of the National Academy of Sciences: 6168–71.
- ¹¹⁷⁶ Kirkpatrick, Mark, Toby Johnson, and Nick Barton. 2002. "General Models of Multilocus Evolution." *Genetics* 161 (4). Genetics: 1727–50.
- ¹¹⁷⁸ 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.
- ¹¹⁸⁰ Kopp, Michael, and Sergey Gavrilets. 2006. "Multilocus Genetics and the Coevolution of Quantitative Traits." *Evolution* 60 (7). Wiley: 1321–36.
- ¹¹⁸² Kölzsch, Andrea, Adriana Alzate, Frederic Bartumeus, Monique de Jager, Ellen J. Weerman, Geerten M. Hengeveld, Marc Naguib, Bart A. Nolet, and Johan van de Koppel. 2015. "Experimental Evidence for Inherent Lévy Search Behaviour in Foraging Animals." *Proceedings of the Royal Society B: Biological Sciences* 282 (1807). The Royal Society: 20150424.
- ¹¹⁸⁶ Kraft, Nathan J. B., William K. Cornwell, Campbell O. Webb, and David D. Ackerly. 2007. "Trait Evolution, Community Assembly, and the Phylogenetic Structure of Ecological Communities." *The American Naturalist* 170 (2). University of Chicago Press: 271–83.
- ¹¹⁸⁹ Krylov, N. V., and B. L. Rozovskii. 1981. "Stochastic Evolution Equations." *Journal of Soviet Mathematics* 16 (4). Springer Science; Business Media LLC: 1233–77.
- ¹¹⁹¹ Lande, Russel. 1976. "Natural Selection and Random Genetic Drift in Phenotypic Evolution." *Evolution* 30 (2). Wiley: 314–34.
- ¹¹⁹³ ———. 1980. "The Genetic Covariance between Characters Maintained by Pleiotropic Mutations." *Genetics* 94 (1): 203–15.
- ¹¹⁹⁵ Lande, Russell. 1982. "A Quantitative Genetic Theory of Life History Evolution." *Ecology* 63 (3). Wiley: 607–15.
- ¹¹⁹⁷ Lande, Russell, and Stevan J. Arnold. 1983. "The Measurement of Selection on Correlated Characters." *Evolution* 37 (6). JSTOR: 1210.
- ¹¹⁹⁹ 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.
- ¹²⁰¹ Landis, Michael J., and Joshua G. Schraiber. 2017. "Pulsed Evolution Shaped Modern Vertebrate Body Sizes." *Proceedings of the National Academy of Sciences* 114 (50). Proceedings of the National Academy of Sciences: 13224–9.
- ¹²⁰⁴ Levins, Richard. 1968. *Evolution in Changing Environments: Some Theoretical Explorations. (MPB-2) (Monographs in Population Biology)*. Princeton University Press.
- ¹²⁰⁶ Li, Zeng-Hu. 1998. "Absolute Continuity of Measure Branching Processes with Interaction." *Chinese Journal of Applied Probability and Statistics* 14. Citeseer: 231–42.
- ¹²⁰⁸ Lion, Sébastien. 2018. "Theoretical Approaches in Evolutionary Ecology: Environmental Feedback as a Unifying Perspective." *The American Naturalist* 191 (1). University of Chicago Press: 21–44.
- ¹²¹⁰ Loreau, Michel. 2010. *From Populations to Ecosystems: Theoretical Foundations for a New Ecological Synthesis*. Princeton University Press.

- 1212 Lynch, Michael, and Bruce Walsh. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates is
1213 an imprint of Oxford University Press.
- 1214 MacArthur, Robert H. 1969. "Species Packing, and what Competition Minimizes." *Proceedings of the National
1215 Academy of Sciences* 64 (4). Proceedings of the National Academy of Sciences: 1369–71.
- 1216 ———. 1970. "Species Packing and Competitive Equilibrium for Many Species." *Theoretical Population
1217 Biology* 1 (1). Elsevier BV: 1–11.
- 1218 ———. 1972. *Geographical Ecology*. Princeton University Press.
- 1219 MacArthur, Robert H., and Richard Levins. 1967. "The Limiting Similarity, Convergence, and Divergence of
1220 Coexisting Species." *The American Naturalist* 101 (921). University of Chicago Press: 377–85.
- 1221 Manceau, Marc, Amaury Lambert, and Hélène Morlon. 2016. "A Unifying Comparative Phylogenetic
1222 Framework Including Traits Coevolving Across Interacting Lineages." *Systematic Biology*, December. Oxford
1223 University Press (OUP), syw115.
- 1224 Marx, Hannah E., Cédric Dentant, Julien Renaud, Romain Delunel, David C. Tank, and Sébastien Lavergne.
1225 2017. "Riders in the Sky (Islands): Using a Mega-Phylogenetic Approach to Understand Plant Species
1226 Distribution and Coexistence at the Altitudinal Limits of Angiosperm Plant Life." *Journal of Biogeography*
1227 44 (11). Wiley: 2618–30.
- 1228 McPeek, Mark A. 2017. *Evolutionary Community Ecology*. Princeton University Press.
- 1229 Meester, Luc De, Kristien I. Brans, Lynn Govaert, Caroline Souffreau, Shinjini Mukherjee, Hélène Vanvelk,
1230 Konrad Korzeniowski, et al. 2018. "Analyzing Eco-Evolutionary Dynamics - the Challenging Complexity of
1231 the Real World." *Functional Ecology*. Wiley.
- 1232 Méléard, M, and S Roelly. 1992. "Interacting Branching Measure Processes." *Stochastic Partial Differential
1233 Equations and Applications* (G. Da Prato and L. Tubaro, Eds.), 246–56.
- 1234 ———. 1993. "Interacting Measure Branching Processes. Some Bounds for the Support." *Stochastics and
1235 Stochastic Reports* 44 (1-2). Informa UK Limited: 103–21.
- 1236 Mubayi, Anuj, Christopher Kribs, Viswanathan Arunachalam, and Carlos Castillo-Chavez. 2019. "Studying
1237 Complexity and Risk Through Stochastic Population Dynamics: Persistence, Resonance, and Extinction in
1238 Ecosystems." In *Handbook of Statistics*, 157–93. Elsevier.
- 1239 Nowak, Martin A. 2006. *Evolutionary Dynamics: Exploring the Equations of Life*. Belknap Press.
- 1240 Nuismer, Scott L., Michael Doebeli, and Danny Browning. 2005. "The Coevolutionary Dynamics of
1241 Antagonistic Interactions Mediated by Quantitative Traits with Evolving Variances." *Evolution* 59 (10). The
1242 Society for the Study of Evolution: 2073.
- 1243 Nuismer, Scott L., and Luke J. Harmon. 2014. "Predicting Rates of Interspecific Interaction from Phylogenetic
1244 Trees." Edited by Jerome Chave. *Ecology Letters* 18 (1). Wiley: 17–27.
- 1245 Nuismer, Scott L., Pedro Jordano, and Jordi Bascompte. 2012. "Coevolution and the Architecture of
1246 Mutualistic Networks." *Evolution* 67 (2). Wiley: 338–54.
- 1247 Nuismer, Scott L., Benjamin J. Ridenhour, and Benjamin P. Oswald. 2007. "Antagonistic Coevolution
1248 Mediated by Phenotypic Differences Between Quantitative Traits." *Evolution* 61 (8). Wiley: 1823–34.
- 1249 Nuismer, Scott L., Bob Week, and Marcelo A. Aizen. 2018. "Coevolution Slows the Disassembly of Mutualistic
1250 Networks." *The American Naturalist* 192 (4). University of Chicago Press: 490–502.
- 1251 Nuland, Michael E. Van, Ian M. Ware, Joseph K. Bailey, and Jennifer A. Schweitzer. 2019. "Ecosystem
1252 Feedbacks Contribute to Geographic Variation in Plant-Soil Eco-Evolutionary Dynamics Across a Fertility
1253 Gradient." Edited by Franziska Brunner. *Functional Ecology* 33 (1). Wiley: 95–106.
- 1254 Page, Karen M., and Martin A. Nowak. 2002. "Unifying Evolutionary Dynamics." *Journal of Theoretical
1255 Biology* 219 (1). Elsevier BV: 93–98.

- 1256 Parachnowitsch, Amy L, Jessamyn S Manson, and Nina Sletvold. 2018. "Evolutionary Ecology of Nectar."
1257 *Annals of Botany* 123 (2). Oxford University Press (OUP): 247–61.
- 1258 Patel, Swati, and Reinhard Bürger. 2019. "Eco-Evolutionary Feedbacks Between Prey Densities and Linkage
1259 Disequilibrium in the Predator Maintain Diversity." *Evolution* 73 (8). Wiley: 1533–48.
- 1260 Perkins, Edwin A. 1991. "Conditional Dawson-Watanabe Processes and Fleming-Viot Processes." In *Seminar
1261 on Stochastic Processes, 1991*, 143–56. Birkhäuser Boston.
- 1262 ———. 1992. "Measure-Valued Branching Diffusions with Spatial Interactions." *Probability Theory and
1263 Related Fields* 94 (2). Springer Science; Business Media LLC: 189–245.
- 1264 ———. 1995. *On the Martingale Problem for Interactive Measure-Valued Branching Diffusions*. Amer
1265 Mathematical Society.
- 1266 Price, George R. 1970. "Selection and Covariance." *Nature* 227 (5257). Springer Nature: 520–21.
- 1267 Queller, David C. 2017. "Fundamental Theorems of Evolution." *The American Naturalist* 189 (4). University
1268 of Chicago Press: 345–53.
- 1269 Reimers, Mark. 1989. "One Dimensional Stochastic Partial Differential Equations and the Branching Measure
1270 Diffusion." *Probability Theory and Related Fields* 81 (3). Springer Nature: 319–40.
- 1271 Robertson, Alan. 1966. "A Mathematical Model of the Culling Process in Dairy Cattle." *Animal Science* 8
1272 (1). Cambridge University Press: 95–108.
- 1273 Roughgarden, Joan. 1979. *Theory of Population Genetics and Evolutionary Ecology: An Introduction*.
1274 Macmillan.
- 1275 Rudman, Seth M., Matthew A. Barbour, Katalin Csilléry, Phillip Gienapp, Frederic Guillaume, Nelson G.
1276 Hairston Jr, Andrew P. Hendry, et al. 2017. "What Genomic Data Can Reveal About Eco-Evolutionary
1277 Dynamics." *Nature Ecology & Evolution* 2 (1). Springer Nature: 9–15.
- 1278 Schreiber, Sebastian J., Swati Patel, and Casey terHorst. 2018. "Evolution as a Coexistence Mechanism:
1279 Does Genetic Architecture Matter?" *The American Naturalist* 191 (3). University of Chicago Press: 407–20.
- 1280 Schuster, Peter, and Karl Sigmund. 1983. "Replicator Dynamics." *Journal of Theoretical Biology* 100 (3).
1281 Elsevier BV: 533–38.
- 1282 Skovmand, Lotte H., Charles C.Y. Xu, Maria R. Servedio, Patrik Nosil, Rowan D.H. Barrett, and Andrew P.
1283 Hendry. 2018. "Keystone Genes." *Trends in Ecology & Evolution* 33 (9). Elsevier BV: 689–700.
- 1284 Sterner, R.W., and J.J. Elser. 2008. "Ecological Stoichiometry: Overview." In *Encyclopedia of Ecology*,
1285 1101–16. Elsevier.
- 1286 Taylor, Peter D., and Leo B. Jonker. 1978. "Evolutionary Stable Strategies and Game Dynamics." *Mathematical Biosciences* 40 (1-2). Elsevier BV: 145–56.
- 1288 Tilman, David. 1982. *Resource Competition and Community Structure*. Princeton University Press.
- 1289 Turelli, Michael. 1984. "Heritable Genetic Variation via Mutation-Selection Balance: Lerchs Zeta Meets the
1290 Abdominal Bristle." *Theoretical Population Biology* 25 (2). Elsevier: 138–93.
- 1291 ———. 1986. "Gaussian Versus Non-Gaussian Genetic Analyses of Polygenic Mutation-Selection Balance." In
1292 *Evolutionary Processes and Theory*, 607–28. Academic Press.
- 1293 ———. 2017. "Commentary: Fisher's Infinitesimal Model: A Story for the Ages." *Theoretical Population
1294 Biology* 118 (December). Elsevier BV: 46–49.
- 1295 Turelli, Michael, and Nick Barton. 1994. "Genetic and statistical analyses of strong selection on polygenic
1296 traits: what, me normal?" *Genetics* 138 (3): 913–41.
- 1297 Walsh, John B. 1986. "An Introduction to Stochastic Partial Differential Equations." In *Lecture Notes in
1298 Mathematics*, 265–439. Springer Berlin Heidelberg.

- ¹²⁹⁹ 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.
- ¹³⁰⁰ Wright, Sewall. 1931. “Evolution in Mendelian Populations.” *Genetics* 16 (2). Genetics: 97–159.
- ¹³⁰¹ Xiao, Xiao, Daniel J. McGinn, 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.