

A White Noise Approach to Evolutionary Ecology

Bob Week, Steve Krone, Scott Nuismer

8/19/2019

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. 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 compute these quantities under the two scenarios of weak and moderate competition and investigate the relationship between selection gradients and competition coefficients. This illustrative investigation suggests magnitudes of selection gradients only weakly correspond to competition coefficients. Future work building on this formalization can contribute to the understanding of feedbacks between evolution and ecology and how random processes such as demographic stochasticity and random genetic drift modulate the strength of these feedbacks.

1 Introduction

A primary aim of theoretical evolutionary ecology is the development of mathematical approaches to describe the evolution of biological 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; McPeck 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 should have 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 take one small step in this direction by establishing a formal connection between the continuous-time dynamics of abundance and quantitative traits in stochastically evolving populations and investigate the relationship between strengths of ecological interactions and coevolution.

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 basis of evolution by natural selection, the mean fitness across all individuals in a population determines the growth, stasis or decline of abundance. In section 2.1 we review the mathematical formalization of this connection, which has been established in the contexts of population genetics (Crow and Kimura 2009), evolutionary game theory (Nowak 2006) and quantitative genetics (Lion 2018). Although reviewing these

accomplishments reveals a beautiful synthesis of evolution and population ecology, it also reveals a gap in theoretical approaches to incorporate the intrinsically random nature of populations. Specifically, in theoretical quantitative genetics the derivation of a populations' response to random genetic drift is derived in discrete time under the assumption of constant population (or effective population) size (Lande 1976). Although this approach conveniently mimics the formalism provided by the Wright-Fisher model of population genetics, real population sizes fluctuate over time. Furthermore, since these fluctuations are themselves stochastic, it seems natural to derive expressions for the evolutionary response to demographic stochasticity and consider how the results relate to characterizations of random genetic drift. This can be done in continuous time for population genetic models without too much technical overhead, assuming a finite number of alleles. However, for populations with a continuum of types, such as a quantitative trait, this becomes a serious mathematical endeavor. Here we close this gap by applying some theory of stochastic partial differential equations (SPDE). Our goal is two-fold: 1) Communicate some useful results of SPDE to as wide of audience as possible. With this goal in mind we will not provide a rigorous treatment of SPDE. Instead, we introduce a set of heuristics that only require the basic concepts of Riemann integration, partial differentiation and some exposure to stochastic ordinary differential equations (SDE). For a concise introduction to SDE, we recommend the primer by Evans (2014). 2) Establish a novel synthetic approach to theoretical evolutionary ecology. To demonstrate the utility of our approach, we derive a model of coevolution in an ecological network and use it to investigate the relationship between competition coefficients and selection gradients.

We begin in section 2.1 with a brief review of progress in modelling the deterministic dynamics of populations. Section 2.2 then introduces a useful set of mathematical tools based on the calculus of white noise. In section 2.3 we outline how to employ these tools to derive a system of SDE modelling abundance dynamics and the evolution of the mean and variance of a quantitative trait. In section 2.4 we derive particular results under the assumption of a normally distributed trait. Section 2.5 extends the previous results under the condition of imperfect inheritance. Section 3 demonstrates the utility of our framework through a derivation of a model of coevolution across a guild of S species participating in exploitative competition along a common resource continuum. In section 3.2 we numerically compute distributions of traits and abundances along with interaction networks parameterized by selection gradients under the scenarios of weak and moderate competition, yielding theoretical possibilities for the distribution of pairwise coevolution in ecological communities. In section 3.3 we investigate the relationship between the strengths of competition and coevolution. Finally, section 4 concludes with a few remarks on the limits of this approach along with future directions to incorporate more explicitly the genetic architecture of populations and feedbacks with ecosystem processes.

2 The framework

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

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

2.1 Deterministic dynamics

The connection between the deterministic dynamics of the abundance of a population and the frequencies of discrete phenotypes within it is well established via heterogeneity in the vital rates of differing phenotypes (ie, demographic heterogeneity, Mubayi et al. 2019). This is particularly evident in the derivation of the evolution of allele frequencies due to natural selection provided by Crow and Kimura (2009, §5.3). Specifically, denoting N the abundance of the population, m_i the growth rate (called the Malthusian parameter in Crow and Kimura 2009) of allele i and p_i the frequency of this allele, we find (using their numbering scheme)

$$\dot{N} = N \sum_i p_i m_i = \bar{m} N, \quad (5.3.3)$$

$$\dot{p}_i = (m_i - \bar{m}) p_i. \quad (5.3.5)$$

Equation (5.3.5) is known in the field of evolutionary game theory as the replicator equation (Nowak 2006; Lion 2018). Instead of being explicitly focused on alleles, the replicator equation describes the fluctuations of relative abundances of various *types* in a population in terms of the vital rates of each type (Taylor and Jonker 1978; Schuster and Sigmund 1983). It is straight-forward to extend the replicator equation to include mutation, which is known as the replicator-mutator equation (Nowak 2006). To derive the replicator-mutator equation one can take the same steps to derive (5.3.5) and inject a model of mutation via a matrix of transition rates (Nowak 2006, eqn. 13.25). Building on these insights, we derive an analog of the replicator-mutator equation for a continuum of types (that is, for a quantitative trait). In particular, for a continuously reproducing population at time t with trait values $x \in \mathbb{R}$ and an abundance density $\nu(x, t)$, the demographic dynamics can be modelled by the general PDE

$$\dot{\nu}(x, t) = F(\nu(x, t), x). \quad (1)$$

More specifically, assuming selection is determined by the growth rate $m(\nu(x, t), x)$, mutation is captured by diffusion with coefficient $\frac{\mu}{2}$, and ignoring all other processes, we set

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

In this case equation (1) becomes a nonlinear reaction-diffusion equation. For brevity we often write the growth rate as just $m(\nu, x)$. As an example, logistic growth combined with quadratic stabilizing selection can be captured using the growth rate $m(\nu, x) = r - \frac{a}{2}(\theta - x)^2 - c \int_{\mathbb{R}} \nu(x, t) dx$ where $a \geq 0$ is the strength of stabilizing selection around the phenotypic optimum $\theta \in \mathbb{R}$, $c \geq 0$ captures the effect of intraspecific competition and $r \in \mathbb{R}$ is the intrinsic growth rate in the absence of abiotic selection. We will always assume $r > 0$. In section 3 we will derive a model of coevolution in which the associated fitness function shares some important properties with this fitness function. In particular, we show in SM §5.1 that fitness functions $m(\nu, x)$ satisfying

$$r - \frac{a}{2}(\theta - x)^2 - c \int_{\mathbb{R}} \nu(x, t) dx \leq m(\nu, x) \leq r - \frac{a}{2}(\theta - x)^2 \quad (3)$$

for some $r, a, c \geq 0$ guarantee finite total abundance $N(t)$, trait mean $\bar{x}(t)$ and trait variance $\sigma^2(t)$ for all $t \geq 0$, so long as these variables have finite initial conditions. Since this is the only form of fitness we consider, for the remainder of this paper we assume

$$N(t) = \int_{\mathbb{R}} \nu(x, t) dx < \infty, \quad (4a)$$

$$|\bar{x}(t)| = \frac{1}{N(t)} \left| \int_{\mathbb{R}} x \nu(x, t) dx \right| < \infty, \quad (4b)$$

$$\sigma^2(t) = \frac{1}{N(t)} \int_{\mathbb{R}} (x - \bar{x}(t))^2 \nu(x, t) dx < \infty, \quad \forall t \geq 0. \quad (4c)$$

124 To derive a replicator-mutator equation from the combination of equations (1) and (2), we employ the chain
 125 rule from elementary calculus. Defining $p(x, t) = \nu(x, t)/N(t)$ as the relative density of trait x at time t , we
 126 have

$$\begin{aligned}\dot{N}(t) &= \frac{d}{dt} \int_{\mathbb{R}} \nu(x, t) dx = \int_{\mathbb{R}} \dot{\nu}(x, t) dx \\ &= \int_{\mathbb{R}} m(\nu, x) \nu(x, t) + \frac{\mu}{2} \frac{\partial^2}{\partial x^2} \nu(x, t) dx \\ &= N(t) \int_{\mathbb{R}} m(\nu, x) p(x, t) + \frac{\mu}{2} \frac{\partial^2}{\partial x^2} p(x, t) dx = \bar{m}(t) N(t) \quad (5)\end{aligned}$$

127 where, for $f : \mathbb{R} \rightarrow \mathbb{R}$, we define, when it exists, $\bar{f}(t) = \int_{\mathbb{R}} f(x) p(x, t) dx$. In particular $\bar{m}(t)$ is the mean
 128 fitness across all trait values in the population. Since $\nu(x, t)$ is continuous in x and $\int_{\mathbb{R}} \nu(x, t) dx < \infty$, the
 129 result $\int_{\mathbb{R}} \frac{\partial^2}{\partial x^2} \nu(x, t) dx = 0$ can be derived from the fundamental theorem of calculus. Biological reasoning
 130 agrees with this result since mutation neither creates nor destroys individuals, but merely changes their type
 131 from their parental type. Taking the same approach, we derive the dynamics of $p(x, t)$ as

$$\begin{aligned}\dot{p}(x, t) &= \frac{\partial}{\partial t} \frac{\nu(x, t)}{N(t)} = \frac{N(t) \dot{\nu}(x, t) - \nu(x, t) \dot{N}(t)}{N^2(t)} \\ &= \frac{1}{N(t)} \left(m(\nu, x) \nu(x, t) + \frac{\mu}{2} \frac{\partial^2}{\partial x^2} \nu(x, t) - \bar{m}(t) \nu(x, t) \right) \\ &= (m(\nu, x) - \bar{m}(t)) p(x, t) + \frac{\mu}{2} \frac{\partial^2}{\partial x^2} p(x, t). \quad (6)\end{aligned}$$

This result closely resembles Kimura's continuum-of-alleles model (Kimura 1965). The primary difference being that our model utilizes diffusion instead of convolution with an arbitrary mutation kernel. Of course, our model of mutation can be derived as an approximation to Kimura's model, which has been referred to as the Gaussian allelic approximation (in reference to the distribution of mutational effects at a given locus of a genome on the values of traits, Bürger 2000), the infinitesimal model (in reference to modelling continuous traits as being encoded by an infinite number of loci each having infinitesimal effect, Barton, Etheridge, and Véber 2017) and the Gaussian descendants approximation (in reference to offspring trait values being normally distributed around their parental values, Turelli 2017). Alternatively, since diffusion is the continuous-time equivalent to convolution against a Gaussian kernel (SM §5.2), our model can also be seen as a special case of the continuum-of-alleles model. Under our model of mutation, the dynamics of the mean trait $\bar{x}(t)$ are derived as

$$\begin{aligned}\dot{\bar{x}}(t) &= \frac{d}{dt} \int_{\mathbb{R}} x p(x, t) dx = \int_{\mathbb{R}} x \dot{p}(x, t) dx \\ &= \int_{\mathbb{R}} x (m(\nu, x) - \bar{m}(t)) p(x, t) + x \frac{\mu}{2} \frac{\partial^2}{\partial x^2} p(x, t) dx \\ &= \text{Cov}_t(x, m(\nu, x)) + \frac{\mu}{2} \int_{\mathbb{R}} x \frac{\partial^2}{\partial x^2} p(x, t) dx, \quad (7)\end{aligned}$$

132 where, for $f, g : \mathbb{R} \rightarrow \mathbb{R}$, we define, when it exists,

$$\text{Cov}_t(f(x), g(x)) = \int_{\mathbb{R}} (f(x) - \bar{f}(t))(g(x) - \bar{g}(t)) p(x, t) dx. \quad (8)$$

133 Equation (7) is a continuous time analog of the well known Robertson-Price equation (Robertson 1966;
 134 Price 1970; Frank 2012; Queller 2017; Lion 2018). In the particular case developed here, transmission

bias only occurs when $p(x, t)$ is not symmetric around \bar{x} . In particular, when $p(x, t)$ is a Gaussian curve $\int_{\mathbb{R}} x \frac{\partial^2}{\partial x^2} p(x, t) dx = 0$. Following this approach, in SM §5.3 we derive the dynamics of the variance $\sigma^2(t) = \text{Cov}_t(x, x)$ as

$$\dot{\sigma}^2(t) = \text{Cov}_t\left((x - \bar{x})^2, m(\nu, x)\right) + \frac{\mu}{2} \int_{\mathbb{R}} (x - \bar{x})^2 \frac{\partial^2}{\partial x^2} p(x, t) dx. \quad (9)$$

In the absence of mutation equation (9) agrees with the result derived by Lion (2018) for discrete phenotypes. When p is Gaussian the increase in variance due to mutation simplifies to μ .

In the following sections we extend these results to the case when PDE (1) is driven by a space-time white noise process $\dot{W}(x, t)$. We begin by defining space-time white noise and illustrating its relevant properties. We then introduce a set of heuristics for performing calculations with respect to $\dot{W}(x, t)$. Since our aim is to present this material to as wide of audience as possible, our approach deviates from standard definitions to remove the need for a detailed technical treatment. For those not interested in the white noise calculus and would rather jump straight into more biologically relevant results, we recommend skipping to §2.4.

2.2 White noise calculus and superprocesses

2.2.1 Definition and basic properties

We define $L_C^2(\mathbb{R} \times [0, \infty))$ as the set of functions $f : \mathbb{R} \times [0, \infty) \rightarrow \mathbb{R}$ that are continuous in their second argument (time) and satisfy $\|f\|_2(t) = \sqrt{\int_{\mathbb{R}} f^2(x, t) dx} < \infty$. 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 $t > 0$ 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 $t, t_1, t_2 \geq 0$,

$$\mathbb{E}(\mathbf{W}_t(f)) = \mathbb{E}(\mathbf{W}_t(g)) = 0, \quad (10)$$

$$\mathbb{C}(\mathbf{W}_{t_1}(f), \mathbf{W}_{t_2}(g)) = \int_0^{t_1 \wedge t_2} \int_{\mathbb{R}} f(x, s) g(x, s) dx ds, \quad (11)$$

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 $t > 0$). 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 $t\|f\|_2^2$. With the generalized process \mathbf{W} defined, we define the space-time white noise $\dot{W}(x, t)$ implicitly via

$$\left\langle \int_0^t \int_{\mathbb{R}} f(x, s) \dot{W}(x, s) dx ds \right\rangle = \mathbf{W}_t(f), \quad \forall f \in L_C^2(\mathbb{R} \times [0, \infty)), \quad t > 0. \quad (12)$$

We place quotations around the left-hand side of the above expression to emphasize its informal nature. Since space-time white noise, represented by the symbol $\dot{W}(x, t)$, does not exist as a function in the classical sense it cannot be integrated as suggested above. However, similar to the approach taken here, space-time white noise can be made sense of in a more abstract setting as a *measure*-valued process (Dawson 1975; Walsh 1986). Hence, our use of integral notation is merely symbolic to provide intuition and should not be confused with classical Riemann integrals. One can think of $\dot{W}(x, t)$ as the static seen on old television sets or infinitely detailed random dust spread across both time and space. Figure 1 illustrates sample paths of white noise in one (left) and two (right) dimensions.

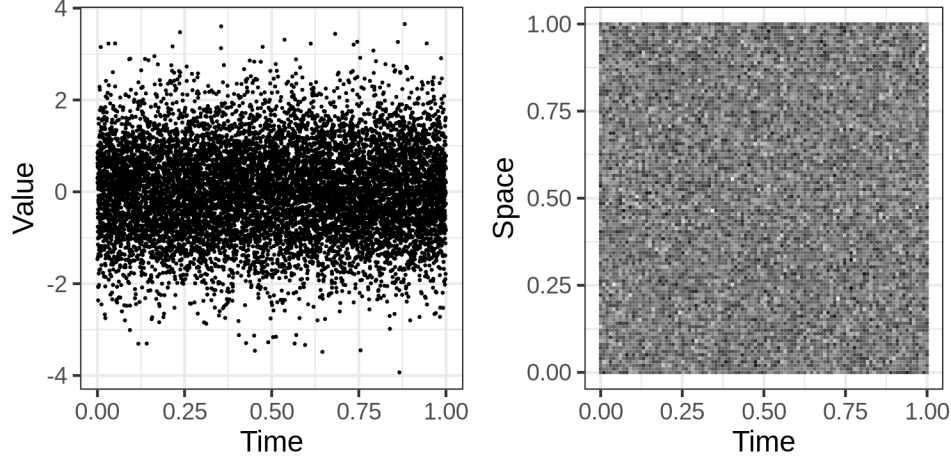


Figure 1: Approximations of sample paths of temporal white noise (left) and space-time white noise (right) with brightness scaled to value.

Following our definition of white noise, when attempting to compute the value of this random field we sample it using $L^2_C(\mathbb{R} \times [0, \infty))$ 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

$$I_{D \times [0, \infty)}(x, t) = \begin{cases} 0, & x \notin D \\ 1, & x \in D \end{cases}. \quad (13)$$

Since $\int_{\mathbb{R}} I_{D \times [0, \infty)}^2(x, t) dx = \int_D dx = |D|$, where $|D|$ denotes the length of D , $I_{D \times [0, \infty)} \in L^2_C(\mathbb{R} \times [0, \infty))$. Thus, by equations (10) and (11), we have

$$\mathbb{E} \left(\int_0^t \int_D \dot{W}(x, s) dx ds \right) = 0, \quad (14)$$

$$\mathbb{V} \left(\int_0^t \int_D \dot{W}(x, s) dx ds \right) = t|D|, \quad (15)$$

where \mathbb{V} denotes the variance operator with respect to the underlying probability space. Following this notation, equations (10) and (11) can be rewritten as

$$\mathbb{E} \left(\int_0^t \int_{\mathbb{R}} f(x, s) \dot{W}(x, s) dx ds \right) = 0, \quad (16)$$

$$\mathbb{C} \left(\int_0^{t_1} \int_{\mathbb{R}} f(x, s) \dot{W}(x, s) dx ds, \int_0^{t_2} \int_{\mathbb{R}} g(x, s) \dot{W}(x, s) dx ds \right) = \int_0^{t_1 \wedge t_2} \int_{\mathbb{R}} f(x, s) g(x, s) dx ds. \quad (17)$$

We write $d\hat{\mathbf{W}}_t(f) = \frac{1}{\|f\|_2(t)} \left(\int_{\mathbb{R}} f(x, t) \dot{W}(x, t) dx \right) dt$ (and typically suppress the dependency on time) so that

$$\int_0^t d\hat{\mathbf{W}}(f) = \int_0^t \int_{\mathbb{R}} \frac{f(x, s)}{\sqrt{\int_{\mathbb{R}} f^2(s, y) dy}} \dot{W}(x, s) dx ds. \quad (18)$$

It follows that $\mathbb{E}(\int_0^t d\hat{\mathbf{W}}(f)) = 0$ and $\mathbb{C}(\int_0^{t_1} d\hat{\mathbf{W}}(f), \int_0^{t_2} d\hat{\mathbf{W}}(f)) = t_1 \wedge t_2$. This implies that, as a function of t , $\int_0^t d\hat{\mathbf{W}}(f)$ is a standard Brownian motion for any $f \in L^2_C(\mathbb{R} \times [0, \infty))$. Hence, $d\hat{\mathbf{W}}(f)$ is analogous

to the traditional shorthand used to denote stochastic differentials. Thus, equation (17) effectively extends Itô's multiplication table to:

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

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

The extension of Itô's multiplication table and properties of white noise outlined in this subsection provide a useful set of tools for working with SPDE. In SM §2.3.1 we employ these tools to derive SDE that track the dynamics of abundance, mean trait and phenotypic variance of a population from a particular SPDE. In the following subsection, we review how this SPDE naturally arises from diffusion limits of branching processes.

2.3 From branching processes to SPDE

In real populations individuals are born and potentially reproduce before they ultimately die. These three events provide the basic ingredients of a branching process. Mathematical investigations of such processes have a relatively deep history (Kendall 1966). The most simple branching process, known as the Galton-Watson process, describes the number of individuals alive at a given time $t \geq 0$ as a non-negative integer (Kimmel and Axelrod 2015). Feller (1951) introduced a formal method to approximate branching processes with diffusion processes which are continuous in state (ie, population size is approximated as a continuous quantity). Since diffusion processes possess greater analytical tractability than branching processes, Feller's method, known as the diffusion limit, has acquired immense popularity, particularly in the field of mathematical population genetics (Ewens 2004). For over the past half of a century a great deal of accomplishments have been achieved in formalizing the diffusion limits of branching processes that describe biological populations with some continuous structure (Watanabe 1968; Dawson 1975; Méléard and Roelly 1993; Perkins 1995; Li 1998; Etheridge 2000, 2008; Bertoin and Le Gall 2003; Barton and Etheridge 2019). This structure can represent geographic space or, relevant to our context, phenotypic space. One particularly important branching process with continuous structure is known as the branching Brownian motion. Unlike other stochastic processes that take values in d -dimensional Euclidean space, branching Brownian motion takes values in the set of non-negative finite measures over \mathbb{R}^d . Intuitively, one can think of a measure as a function that maps bounded subsets of \mathbb{R}^d to real numbers. As a technical aside, we only consider the Borel subsets of \mathbb{R}^d . In particular, denoting X_t a branching Brownian motion, for any $D \subset \mathbb{R}^d$, $X_t(D)$ returns the number of individuals alive within the region D at time $t \geq 0$. The branching Brownian motion has three main components:

- 1) **Spatial movement:** Individuals are born at the current location of their parent and then 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 represents the value of its phenotype.
- 2) **Branching rate:** Lifetimes of these individuals are often assumed to be exponentially distributed. At the death of an individual, a branching event occurs where a random (possibly zero) number of offspring are left. However, in our case we will assume lifetimes are non-random and equal to a single unit of time for all individuals. From a mathematical standpoint this assumption makes little difference since many alternative branching processes lead to the same diffusion limit. However, we make this assumption because we are interpreting movement as mutation and are assuming the effects of mutations are

independent and identically distributed for the trait values inherited by each offspring. Note that a Brownian motion with diffusion parameter $\sqrt{\mu}$ integrated over the unit interval is a zero-mean Gaussian random variable with variance μ . This fact creates a vital link to the deterministic dynamics reviewed above. Indeed, in the absence of selection, the deterministic PDE combining equations (1) and (2) reduces to the Fokker-Planck equation for a Brownian motion with diffusion parameter $\sqrt{\mu}$.

3) **Branching mechanism:** At a branching event, the parent leaves a random number of offspring. The distribution of offspring left is called the branching mechanism. Under the so-called critical condition the average number of offspring left is set equal to one. We denote the variance in reproductive output by V .

To obtain a SPDE from the branching Brownian motion we take a diffusion limit. For a population of size n at time t we rescale the mass of individuals and time by $1/n$, space by $1/\sqrt{n}$ and consider the limit as $n \rightarrow \infty$. That is, denoting $\sqrt{n}D = \{\sqrt{n}x : x \in D \subset \mathbb{R}^d\}$, we consider the rescaled process $X_t^{(n)}(D) = \frac{1}{n}X_{nt}(\sqrt{n}D)$. 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). It turns out that for spatial dimension $d = 1$, \mathcal{X}_t is absolutely continuous for each $t \geq 0$. This means that we can write $\mathcal{X}_t(D) = \int_D \nu(x, t) dx$ for some density process $\nu(x, t)$ (Konno and Shiga 1988; Reimers 1989). Under the critical condition this density process satisfies the SPDE

$$\dot{\nu}(x, t) = \frac{\mu}{2} \Delta \nu(x, t) + \sqrt{V \nu(x, t)} \dot{W}(x, t). \quad (19)$$

The symbol Δ , referred to as the Laplacian or Laplace operator, represents here the continuous extension of the second derivative $\frac{\partial^2}{\partial x^2}$ to the space of bounded continuous functions on \mathbb{R} vanishing at $\pm\infty$. To rigorously interpret SPDE (19), we integrate the solution $\nu(x, t)$ against functions $f \in C_b^\infty(\mathbb{R})$, where $C_b^\infty(\mathbb{R})$ is the set of bounded and infinitely continuously differentiable functions on \mathbb{R} . That is, equation (19) is just an abbreviation for

$$\begin{aligned} \int_{\mathbb{R}} \nu(x, t) f(x) dx - \int_{\mathbb{R}} \nu(x, 0) f(x) dx \\ = \int_0^t \int_{\mathbb{R}} \nu(x, s) \frac{\mu}{2} \Delta f(x) ds dx + \int_0^t \int_{\mathbb{R}} f(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) dx ds. \end{aligned} \quad (20)$$

For more on the general theory of SPDE see Walsh (1986). Note that since $\nu(x, t)$ is the density of a finite measure, it is integrable for each $t \geq 0$. Thus, since $|f(x)| \leq M$ for every x and some $M > 0$, setting $\varphi(x, t) = f(x) \sqrt{V \nu(x, t)}$ implies $\varphi \in L_C^2(\mathbb{R} \times [0, \infty))$. Hence, the white noise integral on the right-hand side of equation (20) can be understood using the heuristics introduced above. Evaluating equation (20) in the particular case of $f(x) \equiv 1$ returns the total mass process, which we refer to as the total abundance $N(t)$.

A convergence theorem for the diffusion limit of a more general version of the branching Brownian motion was established by Watanabe (1968). Dawson (1975) suggested that, for spatial dimension $d = 1$, this diffusion limit should admit a density process that satisfies a SPDE. Konno and Shiga (1988) and Reimers (1989) independently proved that this is the case. The diffusion limit of this more general branching process (in arbitrary spatial dimension) is referred to as the Dawson-Watanabe superprocess (Etheridge 2000). Conditioning the Dawson-Watanabe superprocess to have constant mass returns the Fleming-Viot process (Etheridge and March 1991; Perkins 1991) which has been popular in studies of spatial population genetics. In particular, an extension of the Fleming-Viot process, known as the Λ -Fleming-Viot process, was introduced by Bertoin and Le Gall (2003) and used (and coined) by Etheridge (2008) to resolve some technical challenges in modelling isolation by distance (Felsenstein 1975; see also Barton, Etheridge, and Véber 2013; and Barton and Etheridge 2019).

A more general version of the Dawson-Watanabe superprocess that includes interactions (which can represent, among other phenomena, density-dependent growth rates and frequency-dependent selection) was studied by Perkins (1992, 1995). When the interactions occur only at branching events, which in our case translates

to independence of mutation and selection, the Dawson-Watanabe superprocess with interactions on one dimensional trait space has also been shown to satisfy a SPDE that admits function-valued solutions which are positive, integrable and continuous in time and space (Li 1998). Hence, this process provides a very natural and mathematically convenient candidate to base our framework on.

Instead of formally developing the diffusion limit of a branching Brownian motion with interaction we skip directly to a special case of the SPDE found by Li (1998) mentioned above:

$$\dot{\nu}(x, t) = m(\nu, x)\nu(x, t) + \frac{\mu}{2}\Delta\nu(x, t) + \sqrt{V\nu(x, t)}\dot{W}(x, t). \quad (21)$$

Note that, comparing equation (21) to equation (3.5) of Li (1998), our m and V correspond to Li's b and c respectively. Denoting $C_1(\mathbb{R})^+$ the set of non-negative, continuous, integrable functions on \mathbb{R} , it is important to note equation (21) is only formal when $m(h, x)$ is bounded across all $h \in C_1(\mathbb{R})^+$ and $x \in \mathbb{R}$. However, recalling condition (3), the $m(h, x)$ we consider are only bounded above. Replacing $m(h, x)$ with its upper-bound demonstrates that the associated superprocess should still take values in the set of finite measures. Hence, we should still be able to compute the total mass process $N(t)$. What remains is to show the spatial moments $\int_{\mathbb{R}} |x|^n \nu(x, t) dx$ are finite for each $t > 0$ and $n \in \mathbb{N}$, assuming finite initial conditions. In SM §5.4 we investigate this claim by comparing simulations of the Branching brownian motion with selection to analytical solutions of the PDE obtained by combining equations (1) and (2). Our results suggest this assertion is sound, but it has yet to be rigorously justified. For now, we leave this as an open problem and assume all spatial moments are finite for every $t \geq 0$.

2.3.1 Derivation of SDE for N , \bar{x} and σ^2

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

$$\begin{aligned} \int_{\mathbb{R}} f(x) \nu(x, t) dx - \int_{\mathbb{R}} f(x) \nu(x, 0) dx \\ = \int_0^t \int_{\mathbb{R}} \nu(x, s) \left(m(\nu, x) f(x) + \frac{\mu}{2} \Delta f(x) \right) dx ds + \int_0^t \int_{\mathbb{R}} f(x) \sqrt{V\nu(x, s)} \dot{W}(x, s) dx ds \end{aligned} \quad (22)$$

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

$$N(t) - N(0) = \bar{m}(t)N(t) + \int_0^t \int_{\mathbb{R}} \sqrt{V\nu(x, s)} \dot{W}(x, s) dx ds. \quad (23)$$

Equation (17) implies

$$\mathbb{V} \left(\int_0^t \int_{\mathbb{R}} \sqrt{V\nu(x, s)} \dot{W}(x, s) dx ds \right) = \int_0^t \int_{\mathbb{R}} V\nu(x, s) dx ds = \int_0^t VN(s) ds. \quad (24)$$

Following the notation introduced in §2.2.1, we have

$$\int_0^t d\hat{\mathbf{W}}_s(\sqrt{V\nu(x, s)}) = \int_0^t \int_{\mathbb{R}} \frac{\sqrt{V\nu(x, s)}}{\sqrt{VN(s)}} \dot{W}(x, s) dx ds \quad (25)$$

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

$$dN(t) = \bar{m}(t)N(t)dt + \sqrt{VN(t)}dW_1(t). \quad (26)$$

Under the assumption that $\int_{\mathbb{R}}(|x| + x^2)\nu(x, t)dx < \infty$ for each $t \geq 0$, we show in SM §5.5 SDE for \bar{x} and σ^2 can be expressed as

$$d\bar{x}(t) = \text{Cov}_t(x, m(\nu, x))dt + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x, t)dx + \sqrt{V \frac{\sigma^2(t)}{N(t)}}dW_2(t), \quad (27)$$

$$d\sigma^2(t) = \left(\text{Cov}_t((x - \bar{x}(t))^2, m(\nu, x)) + \frac{\mu}{2} \int_{\mathbb{R}} (x - \bar{x}(t))^2 \Delta p(x, t)dx - V \frac{\sigma^2(t)}{N(t)} \right) dt + \sqrt{V \frac{(x - \bar{x}(t))^4 - \sigma^4(t)}{N(t)}}dW_3(t). \quad (28)$$

where $p(x, t) = \nu(x, t)/N(t)$ and W_2, W_3 are standard Brownian motions. In SM §5.5 we show that in general W_1 is independent of both W_2 and W_3 , but W_2 and W_3 covary. The expressions derived can be used to investigate the dynamics of the mean and variance for general ν . However, in the next section we further simplify these expressions by approximating ν with a Gaussian curve. In SM §5.5 we show that under the Gaussian case W_1, W_2 and W_3 are independent.

Before we make the Gaussian assumption, there is still quite a bit we can learn from expressions (27) and (28). Firstly, setting $V = 0$ recovers the deterministic dynamics derived in §2.1. Relating the response to demographic stochasticity derived here to the effect of random genetic drift derived in classic quantitative genetic theory, if we pretend that $\sigma^2(t)$ and $N(t)$ are constant with respect to time, then integrating the stochastic term in equation (27) over a single unit of time returns a normally distributed random variable with mean zero and variance equal to $V\sigma^2/N$. In particular, assuming perfect inheritance, when $V = 1$ this random variable coincides with the effect of random genetic drift on the change in mean trait over a single generation derived using sampling arguments (Lande 1976). There is also an interesting connection with classical population genetics. A fundamental result from early population genetic theory is the expected reduction in diversity due to the chance loss of alleles in finite populations (Fisher 1923; Wright 1931). This expected reduction in diversity due to random genetic drift is captured by the third term in the deterministic component of expression (28), but is now parameterized by the variance in reproductive output V .

2.4 Particular results assuming a Gaussian phenotypic distribution

Gaussian phenotypic distributions are often obtained through Gaussian, exponential or weak selection approximations together with a simplified model of inheritance and random mating (Lande 1980; Turelli 1984, 1986, 2017; Bürger 2000). Alternatively, it has been shown that a Gaussian distribution can provide a reasonable approximation even when selection is strong and non-Gaussian (Turelli and Barton 1994). However, our approach adds an additional layer of difficulty. Even with Gaussian selection, the resulting solution to SPDE (21) will only be Gaussian in expectation, assuming a Gaussian initial condition. Yet this difficulty is not as challenging as it may first appear. Indeed, since SPDE (21) can be derived as a diffusion limit we know that, under the appropriate assumptions on selection, genetic architecture and reproduction, the stochastic departure from a Gaussian curve is negligible when the ratio V/N is small (ie, when the variance in reproductive output is much smaller than the population size). In SM §5.4 we demonstrate this result using numerical methods. Mathematically, this requirement restricts model parameters to regions that maintain large population sizes. Biologically, this implies populations are not at risk of extinction. Hence, models developed in this framework are not suitable for studying colonization-extinction dynamics or evolutionary rescue. Allowing for these restrictions, we may safely assume that ν is approximately Gaussian and justify writing

$$\nu(x, t) = \frac{N(t)}{\sqrt{2\pi\sigma^2(t)}} \exp\left(-\frac{(x - \bar{x}(t))^2}{2\sigma^2(t)}\right). \quad (29)$$

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

$$\text{Cov}(x, m) = \sigma^2 \left(\frac{\partial \bar{m}}{\partial \bar{x}} - \frac{\partial \overline{m}}{\partial \bar{x}} \right). \quad (30)$$

Under the same assumption, in SM §5.3 we show

$$\text{Cov}\left((x - \bar{x})^2, m\right) = 2\sigma^4 \left(\frac{\partial \bar{m}}{\partial \sigma^2} - \frac{\partial \overline{m}}{\partial \sigma^2} \right) \quad (31)$$

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

$$d\bar{x} = \sigma^2 \left(\frac{\partial \bar{m}}{\partial \bar{x}} - \frac{\partial \overline{m}}{\partial \bar{x}} \right) dt + \sqrt{V \frac{\sigma^2}{N}} dW_2, \quad (32)$$

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

Note that in the above expressions, the partial derivatives of \bar{m} represent frequency independent selection and the averaged partial derivatives of m represent frequency dependent selection. This relationship has already been pointed out by Lande (1976) for the evolution of trait means, but here we see this relationship holds also for the evolution of trait variance. In the next subsection we generalize this result to the case when traits are imperfectly inherited. In this case, the phenotypic variance σ^2 is replaced by a genetic variance G that represents the proportion of σ^2 explained by additive interactions of alleles at the same locus, but on different chromosomes, as well as the additive interactions among loci encoding for the focal phenotype (Roughgarden 1979; Bulmer 1980; Lynch and Walsh 1998). For completeness we mention that any non-additive effect of interacting alleles at the same locus, but on different chromosomes, is referred to as dominance. Similarly, any non-additive effects among loci encoding for a quantitative character is referred to as epistasis. The variance G is therefore fittingly referred to as the additive genetic variance.

2.5 The evolution of additive genetic variance

To model imperfect heritability we consider the relationship between the expressed phenotype x and its genetic component $g \in \mathbb{R}$ known as the *breeding value* (Lynch and Walsh 1998). We assume the absence of dominance and epistasis. The additive genetic variance can be modelled as the variance of breeding values in a population (Bulmer 1980). For any given individual with breeding value g , we assume the expressed phenotype is normally distributed around g with variance η . We denote the density of this distribution by ψ so that

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

In terms of classic quantitative genetics η is known as the *variance of environmental deviation* (Lynch and Walsh 1998) and ψ plays a similar role as the *segregation kernel* (Roughgarden 1979). To include this relationship in our framework, we write $\rho(g, t)$ as the abundance density of breeding values at time t . Hence $\int_{\mathbb{R}} \rho(g, t) dg = \int_{\mathbb{R}} \nu(x, t) dx = N(t)$. We switch our focus from directly modelling the evolution of $\nu(x, t)$ to modelling the evolution of $\rho(g, t)$. Once $\rho(g, t)$ is defined, we can compute $\nu(x, t)$ via

$$\nu(x, t) = \int_{\mathbb{R}} \rho(g, t) \psi(x|g) dg. \quad (35)$$

However, since selection acts on expressed phenotypes, we have to use our assumed relationship between breeding values and expressed traits to calculate the fitness of breeding values. Using equation (35) we can write the Malthusian fitness of trait value x as

$$m(\nu(x, t), x) = m\left(\int_{\mathbb{R}} \rho(g, t) \psi(x|g) dg, x\right) = \tilde{m}(\rho, x). \quad (36)$$

Hence, we can also calculate the net Malthusian fitness of a breeding value g as

$$m^*(\rho, g) = \int_{\mathbb{R}} \tilde{m}(\rho, x) \psi(x|g) dx. \quad (37)$$

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

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

If we assume $\rho(g, t)$ is Gaussian its mode coincides with \bar{x} and, denoting G the variance of breeding values (ie, the additive genetic variance), $\sigma^2 = G + \eta$. Hence, narrow sense heritability can be computed as $h^2 = G/(G + \eta)$. Furthermore, since

$$\begin{aligned} \bar{m}(t) &= \int_{\mathbb{R}} m(\nu, x) p(x, t) dx = \frac{1}{N(t)} \int_{\mathbb{R}} \tilde{m}(\rho, x) \int_{\mathbb{R}} \rho(g, t) \psi(x|g) dg dx \\ &= \frac{1}{N(t)} \int_{\mathbb{R}} \rho(g, t) \int_{\mathbb{R}} \tilde{m}(\rho, x) \psi(x|g) dx dg = \frac{1}{N(t)} \int_{\mathbb{R}} m^*(\rho, g) \rho(g, t) dg = \bar{m}^*(t), \end{aligned} \quad (39)$$

equations (32) and (33) become

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

$$dG = 2G^2 \left(\frac{\partial \bar{m}}{\partial G} - \frac{\partial \bar{m}}{\partial G} \right) dt + \left(\mu - V \frac{G}{N} \right) dt + G \sqrt{\frac{2V}{N}} dW_3. \quad (41)$$

3 A model of diffuse coevolution

3.1 Formulation

In this section we demonstrate the use of our framework by formulating a model of diffuse coevolution across a guild of S species whose interactions are mediated by resource competition along a single niche axis. Our approach mirrors closely the theory developed by MacArthur and Levins (1967), Levins (1968) and MacArthur (1969, 1970, 1972). The most significant difference, aside from allowing evolution to occur, is the treatment of resource quality, which we replace with a model of abiotic stabilizing selection. A derivation is provided in SM §5.6.

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

$$u_i(\zeta, x_i) = \frac{U_i}{\sqrt{2\pi w_i}} \exp\left(-\frac{(x_i - \zeta)^2}{2w_i}\right). \quad (42)$$

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

$$e_i(\zeta) = Q_i \exp\left(-\frac{A_i}{2}(\theta_i - \zeta)^2\right). \quad (43)$$

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

$$\int_{\mathbb{R}} e_i(\zeta) u_i(\zeta, x_i) d\zeta = \frac{Q_i U_i}{\sqrt{A_i w_i + 1}} \exp\left(-\frac{A_i}{2(A_i w_i + 1)}(\theta_i - x_i)^2\right). \quad (44)$$

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

$$\mathcal{O}_{ij}(x_i, x_j) = \int_{\mathbb{R}} u_i(\zeta, x_i) u_j(\zeta, x_j) d\zeta = \frac{U_i U_j}{\sqrt{2\pi(w_i + w_j)}} \exp\left(-\frac{(x_i - x_j)^2}{2(w_i + w_j)}\right). \quad (45)$$

In SM §5.6 we combine this niche model with equations (26), (40) and (41) to find

$$dN_i = \left\{ R_i - \frac{a_i}{2}((\bar{x}_i - \theta_i)^2 + G_i + \eta_i) + c_i \left(\sqrt{\frac{b_{ii}}{2\pi}} - \sum_{j=1}^S N_j \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2} \right) \right\} N_i dt + \sqrt{V_i N_i} dW_1, \quad (46a)$$

$$d\bar{x}_i = \left\{ a_i G_i (\theta_i - \bar{x}_i) - c_i G_i \left(\sum_{j=1}^S N_j b_{ij} (\bar{x}_j - \bar{x}_i) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2} \right) \right\} dt + \sqrt{V_i \frac{G_i}{N_i}} dW_2, \quad (46b)$$

$$dG_i = \left\{ \mu_i - a_i G_i^2 + c_i G_i^2 \left(\sum_{j=1}^S N_j b_{ij} (1 - b_{ij}(\bar{x}_i - \bar{x}_j)^2) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2} \right) - V_i \frac{G_i}{N_i} \right\} dt + G_i \sqrt{\frac{2V_i}{N_i}} dW_3, \quad (46c)$$

where

$$R_i = \ln \left(\frac{Q_i U_i}{\sqrt{1 + A_i w_i}} \right), \quad (47a)$$

$$a_i = \frac{A_i}{1 + A_i w_i}, \quad (47b)$$

$$b_{ij}(t) = b_{ji}(t) = (w_i + w_j + \eta_i + \eta_j + G_i(t) + G_j(t))^{-1}, \quad (47c)$$

$$c_i \geq 0. \quad (47d)$$

3.2 Community dynamics

For the sake of illustration we numerically integrated system (46) for a richness of $S = 50$ species under two scenarios. For the first scenario of weak competition we randomly chose $c_i \sim \text{Exp}(0.01)$ for each $i = 1, \dots, S$ and for the second scenario of moderately strong competition we randomly chose $c_i \sim \text{Exp}(0.1)$ for each $i = 1, \dots, S$. Background parameters are drawn independently for each species from a common distribution. Specifically, we set model parameters according to Table 2.

Table 2: UNDER CONSTRUCTION Distributions of model parameters under the scenarios of weak and moderate competition. The index i runs through all $1, \dots, S$ species.

Parameter	Description	Weak Competition	Moderate Competition
R_i	innate growth rate, see §3.3	Unif(0.0, 1.0)	same
θ_i	abiotic optimum	Unif(-1.0, 1.0)	same
a_i	strength of abiotic selection	Exp(0.01)	same
c_i	strength of competition	Exp(0.01)	Exp(0.1)
w_i	niche breadth	Unif(0.01, 1.0)	same
U_i	total niche use	Unif(0.01, 1.0)	same
η_i	segregation variance	Unif(1.0e-5, 1.0e-3)	same
μ_i	mutation rate	Exp(0.01)	same
V_i	variance of reproductive output	Exp(0.01)	same

With these two sets of model parameters, we simulated our model for 1.0e3 units of time. For both scenarios, we initialized the abundances to $N_i = 5.0e3$, trait means to $\bar{x}_i = 0.0$ and heritability to $h_i^2 = 0.5$ for each $i = 1, \dots, S$. Temporal dynamics for each scenario are provided in Figure 2. Interaction networks and the distribution of coevolution at $t = 1e3$ for each scenario are displayed in Figure 3.

3.3 The relation between the strength of ecological interactions and coevolution

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

$$\alpha_{ij}(t) = \frac{c_i}{r_i(t)} \int_{\mathbb{R}} \int_{\mathbb{R}} p_i(x, t) p_j(y, t) \mathcal{O}_{ij}(x, y) dx dy = \frac{c_i}{r_i(t)} \sqrt{\frac{b_{ij}(t)}{2\pi}} \exp \left(-\frac{b_{ij}(t)}{2} (\bar{x}_i(t) - \bar{x}_j(t))^2 \right), \quad (48)$$

where

$$r_i(t) = R_i - \frac{a_i}{2} ((\bar{x}_i(t) - \theta_i)^2 + G_i(t) + \eta_i) + c_i \sqrt{\frac{b_{ii}(t)}{2\pi}}. \quad (49)$$

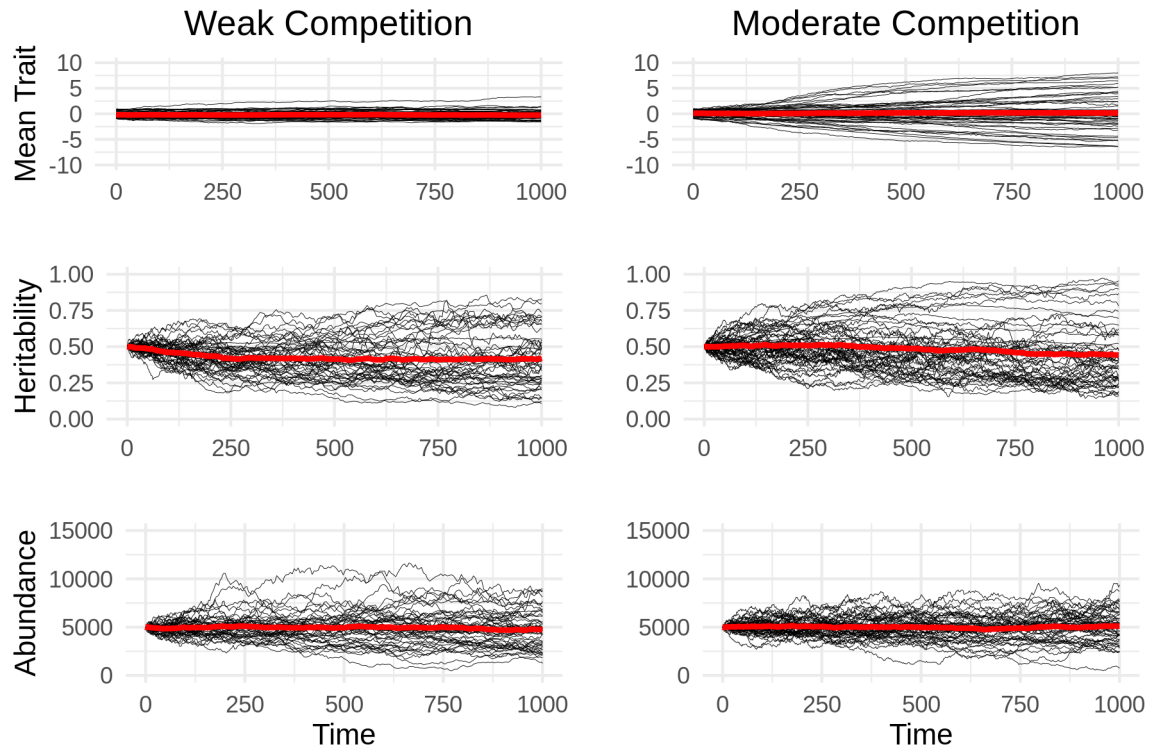


Figure 2: Temporal dynamics of mean trait (top), heritability (middle) and abundance (bottom) for the scenario of weak competition (left) and strong competition (right). Red lines indicate average trend across species.

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

$$dN_i(t) = r_i(t) \left(1 - \sum_{j=1}^S \alpha_{ij}(t) N_j(t) \right) N_i(t) dt + \sqrt{V_i N_i(t)} dW_1(t). \quad (50)$$

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

In SM §5.6 we show that the standardized directional selection gradient (sensu Lande and Arnold 1983) induced by species j on species i can be computed as

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

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

Below we compare interaction networks parameterized by α_{ij} with those parameterized by standardized selection gradients to investigate the correspondence of interaction intensity and coevolutionary change. However, we can already identify one major discrepancy; α_{ij} is maximized when $\bar{x}_i = \bar{x}_j$, but $\beta_{ij} = 0$ under the same condition. We therefore include in our metric of selection the standardized stabilizing selection gradient γ which measures the effect of stabilizing or disruptive selection on phenotypic variance (Lande and Arnold 1983). In SM §5.6 we show that the standardized stabilizing selection gradient induced by species j on species i can be computed as

$$\gamma_{ij}(t) = c_i N_j(t) b_{ij}(t) \left(1 - b_{ij}(t) (\bar{x}_i(t) - \bar{x}_j(t))^2 \right) \sqrt{\frac{b_{ij}(t)}{2\pi}} \exp \left(-\frac{b_{ij}(t)}{2} (\bar{x}_i(t) - \bar{x}_j(t))^2 \right). \quad (52)$$

Note that when $b_{ij}^{-1} > (\bar{x}_i - \bar{x}_j)^2$, $\gamma_{ij} > 0$ and when $b_{ij}^{-1} < (\bar{x}_i - \bar{x}_j)^2$, $\gamma_{ij} < 0$. To measure the total evolutionary change in species i induced by species j , we form the metric $\Psi_{ij} = |\beta_{ij}| + |\gamma_{ij}|$. The top row of Figure 3 displays interaction networks under weak and moderate competition where the edge width connecting species i and j is proportional to $\alpha_{ij}\alpha_{ji}$. The bottom row of Figure 3 displays the distributions of pairwise coevolutionary change, which we measure via $\Psi_{ij}\Psi_{ji}$, under weak and moderate competition.

We now make use of the expressions derived for competition coefficients and selection gradients to investigate their relationship. As a first pass, let us assume all model parameters are equivalent across species and that each species has the same abundance and trait variance. Let us further assume that richness S is large and the distribution of mean trait values is normal with mean $\bar{\bar{x}}$, variance $V_{\bar{\bar{x}}}$ and density $f_{\bar{\bar{x}}}$. Such assumptions are typical when deriving analytical results in the field of theoretical coevolutionary community ecology (Nuismer, Jordano, and Bascompte 2012; Nuismer, Week, and Aizen 2018). If $V_{\bar{\bar{x}}}$ is not too large, then we may approximate r_i with

$$\bar{r} = \int_{\mathbb{R}} \left(R - \frac{a}{2} ((\bar{x} - \theta)^2 + G + \eta) + c \sqrt{\frac{b}{2\pi}} \right) f_{\bar{\bar{x}}}(\bar{x}) d\bar{x} = R - \frac{a}{2} ((\bar{\bar{x}} - \theta)^2 + V_{\bar{\bar{x}}} + G + \eta) + c \sqrt{\frac{b}{2\pi}} \quad (53)$$

Write $\alpha(\bar{x}_i, \bar{x}_j) = \alpha_{ij}$, $\beta(\bar{x}_i, \bar{x}_j) = \beta_{ij}$ and $\gamma(\bar{x}_i, \bar{x}_j) = \gamma_{ij}$. In SM §5.7 we use these assumptions to calculate the first and second order moments describing the joint distribution of competition coefficients and selection gradients across the community. In particular, the correlations between these quantities are found to be:

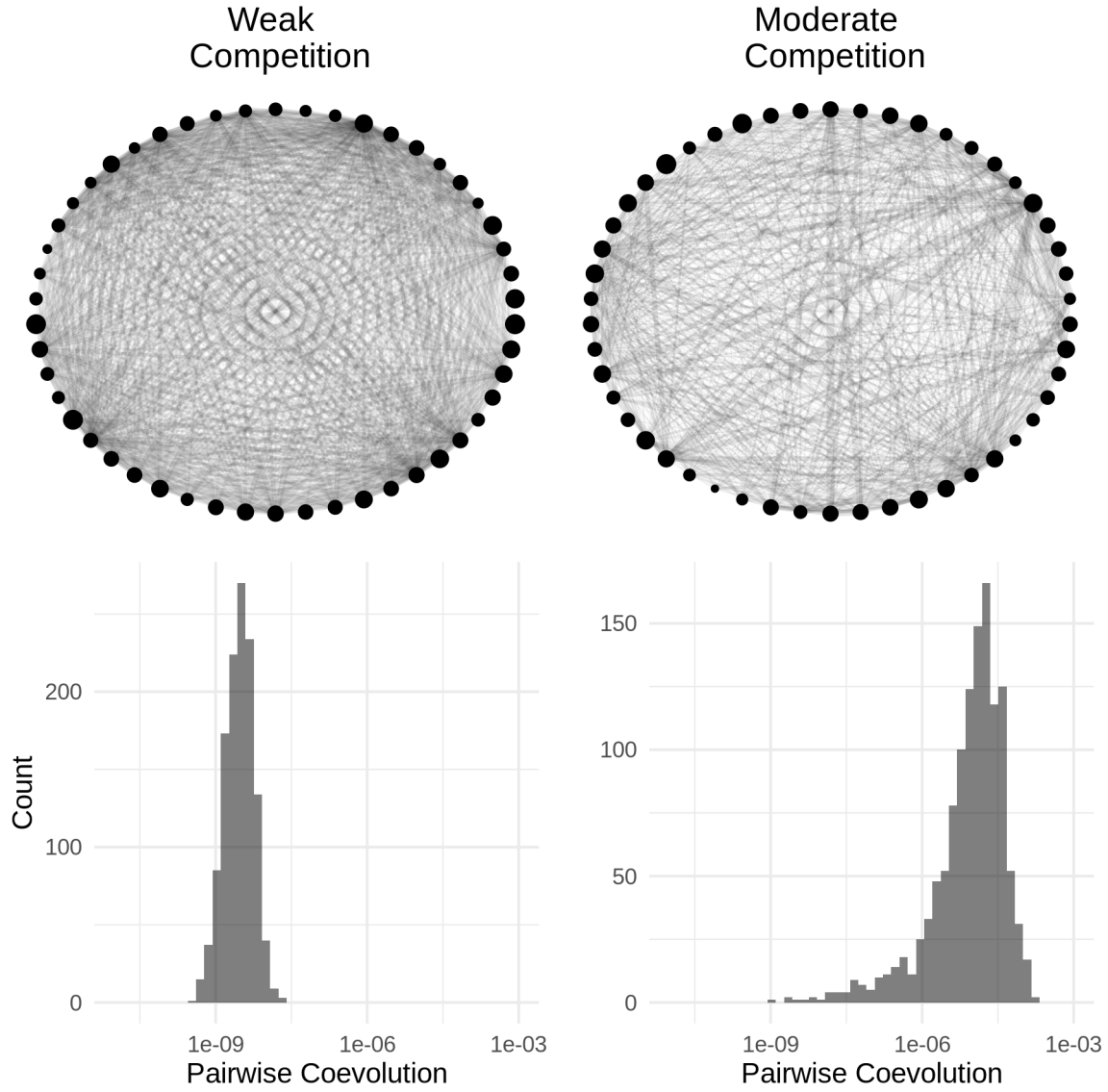


Figure 3: UNDER CONSTRUCTION Networks of interspecific interactions (top row) and distributions of pairwise coevolution (bottom row) under the scenarios of weak competition (left) and moderate competition (right) at time $t = 1.0e3$. Node sizes are proportional to population sizes. Edge widths and shade are monotonically increasing functions of pairwise coevolution.

$$\text{Corr}_{f_{\bar{x}}}(\alpha, \beta) = \quad (54a)$$

$$\text{Corr}_{f_{\bar{x}}}(\alpha, \gamma) = \quad (54b)$$

$$\text{Corr}_{f_{\bar{x}}}(\gamma, \beta) = \quad (54c)$$

4 Conclusion

We have introduced an approach to derive models of evolutionary ecology using the calculus of white noise, demonstrated our approach by deriving a model of diffuse coevolution and investigated the relationship between coevolutionary change and competition coefficients, finding that these quantities are weakly correlated with each other. Although our approach has the merit of rigorously synthesizing the dynamics of abundance and phenotypic evolution of populations, there remains nuance details and important limitations that have yet to be confronted for gaining a more thorough and rigorous understanding of ecological communities. We touch on just five of them here.

Limitations of diffusion limits

As noted by Feller (1951), although diffusion limits provide reasonable approximations for large populations, relatively small populations require discrete models. Hence, as a diffusion limit, SPDE (21) restricts parameter values to regions that maintain large population sizes. This puts an important restriction on any model developed under this framework by implying populations are not at risk of extinction. Hence, studies of evolutionary rescue and colonization-extinction dynamics must be pursued via a different approach. A natural alternative can be developed utilizing the underlying branching Brownian motion process that SPDE (21) is a diffusion limit of (see section 2.3). This process explicitly tracks individuals throughout their life-history and captures reproduction as branching events and therefore models population size n as a non-negative integer instead of a continuously varying positive number. In particular, there is no restriction that prevents $n = 0$. We avoided use of this process here because the resulting models are less analytically tractable and numerical implementations become computationally expensive for large n .

The genetic architecture and distributions of traits

Our model of coevolution rests on the assumption of normally distributed traits, which itself depends on a simplistic model of genetic architecture (see sections 2.1 and 2.5). However, real traits are not encoded by an infinite number of loci each contributing an additive infinitesimal effect and mutations are not inherited as normally distributed deviations from parental breeding values. Departures from this model of genetic architecture can produce more complex distributions of breeding values and expressed traits. Such deviations can be reinforced via strong non-Gaussian selection surfaces, including the surface $m(\nu, x)$ we have derived from niche theory, along with non-random mating. Gaussian approximations are convenient since they are defined by their mean and variance. Hence, future work investigating the effects of non-normally distributed traits on the structure and dynamics of ecological communities will need to confront higher moments such as skew and kurtosis, ideally integrating previously established mathematical approaches to derive the dynamics of these higher moments (Débarre, Yeaman, and Guillaume 2015).

An alternative approach to breaking the assumption of normally distributed trait values is the development of explicit multilocus models. These models describe the contributions of alleles at particular loci in the genome to the development of quantitative traits. Tracking the fluctuations of allele frequencies then allows theoreticians to investigate the dynamics of phenotypic distributions that deviate from normality. While this approach has a long history in theoretical quantitative genetics (Bulmer 1980; Turelli and Barton 1994; Kirkpatrick, Johnson, and Barton 2002) and coevolutionary theory (Nuismer, Doebeli, and Browning 2005; Kopp and Gavrillets 2006; Nuismer, Ridenhour, and Oswald 2007), work to investigate the role of genetic

architecture in mediating feedbacks between the dynamics of population abundances and the distributions of traits mediating ecological interactions has only just begun (Patel and Bürger 2019).

The distinction of niche overlap and competition

Past work has demonstrated several limitations in measuring competition via niche overlap (Abrams 1980; Holt 1987). In particular, Abrams (1980) argues the MacArthur-Levins framework utilized here is unacceptable when nothing is known about the biological mechanism mediating competition. Hence, we must stress that our approach should not be blindly applied to arbitrary communities, but that a great deal of work interpreting our model with respect to the system is a prerequisite task. For example, in the case of a sub-alpine flowering plant community, the niche axis can be interpreted as day-of-year (under a suitable transformation). Our model then assumes within-season fluctuations in the ensemble of abiotic factors including temperature, precipitation, soil nutrients and pollinator activity induces an emergent abiotic optimal date that maximizes fitness in the absence of competition. This optimum likely differs across species with significantly different physiologies and potentially carries a phylogenetic signal. We therefore expect $\theta_i \neq \theta_j$ for our interpretation with respect to a sub-alpine flower community. Individual plants competing for their optimal date to flower then induce the exploitative competition we are modelling.

Another component of Abrams' critique is the distinction between niche overlap and measurements of competition. We also agree that this is an important distinction and is explicitly captured in our model by decomposing the effect of competition on fitness for individuals in species i as the product of niche overlap and a number $c_i > 0$ which we have dubbed the strength of competition. In our model, niche overlap merely determines whether or not an interaction occurs and to what frequency. The strength of competition then determines what effect this interaction has on fitness. A natural generalization is to assume the strength of competition between species i and j can be captured by a function of some other trait values. For example, if y_i represents some other trait in species i , potentially morphological (flower shape) or physiological (water use efficiency), determining the outcome of indirect interspecific interactions between competitors (pollinator preference or nutrient availability) we can model the strength of competition between species i and j as a function f of the distribution of this trait in each species. If this trait correlates with niche centers across individuals within species, or if resource utilization curves u_i depend on y_i , or if f depends on niche centers, the effect on fitness due to competition can no longer be decomposed multiplicatively into distinct components. Hence, our model also requires the necessary assumptions for this decomposition to hold, which vary from system to system.

In another critique by Holt (1987), it is argued that utilizations of qualitatively distinct resources may have non-multiplicative effects on components of fitness. This consideration has also been incorporated into our model. In our derivation of fitness in SM §5.6, we show that distinct components of the growth rate of populations can be decomposed additively under certain circumstances, though further work is needed to develop a general roadmap for when components of fitness can be decomposed additively or multiplicatively.

The role of ecological stoichiometry

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

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

Accounting for macroevolutionary history

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

Of course microevolutionary and ecological dynamics are not completely irrelevant for understanding patterns in communities that are primarily structured by deep evolutionary processes. In particular, macroevolutionary trait evolution is simply the aggregation of microevolutionary change occurring over large spans of time. This suggests a road forward to connect the theory we have introduced to models of macroevolutionary trait evolution that differs from previous approaches (Manceau, Lambert, and Morlon 2016). These previous approaches make the implicit assumption that trait evolution is scale free (ie, self-similar) so that the stochastic evolution of traits on macroevolutionary time scales has the same properties of trait evolution on microevolutionary time scales. In fact, many of these models simply repurpose microevolutionary models by blindly rescaling time from the units of generations to millions of years (Nuismer and Harmon 2014). Although some stochastic processes including Brownian motion do exhibit self-similarity, others do not. For example, consider a modification of the Ornstein-Uhlenbeck process defined by the SDE

$$dX_t = a(\theta_t - X_t)dt + b dW_t \quad (55)$$

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

Final remarks

Although top-down approaches to community ecology, such as the Maximum Entropy Theory of Ecology (Harte 2011), have enjoyed some success in describing community-level patterns (Harte and Newman 2014;

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

5 Supplementary material (SM)

5.1 Sufficient conditions for finite mean, variance and total abundance

We assume, for $m : C_1(\mathbb{R})^+ \times \mathbb{R} \rightarrow \mathbb{R}$, there exists $\theta \in \mathbb{R}$ and $a, c, r > 0$ such that, for each $(f, x) \in C_1(\mathbb{R})^+ \times \mathbb{R}$,

$$r - \frac{a}{2}(\theta - x)^2 - c \int_{\mathbb{R}} f(x) dx \leq m(f, x) \leq r - \frac{a}{2}(\theta - x)^2. \quad (56)$$

Assume that $\nu_0 \in C_1(\mathbb{R})^+$ is twice continuously differentiable and satisfies

$$\int_{\mathbb{R}} |x| \nu_0(x) dx, \int_{\mathbb{R}} x^2 \nu_0(x) dx < \infty. \quad (57)$$

Then

$$\begin{aligned} \int_{\mathbb{R}} |m(\nu_0, x)| \nu_0(x) dx &\leq \int_{\mathbb{R}} \left(r + \frac{a}{2} x^2 + c \int_{\mathbb{R}} \nu_0(y) dy \right) \nu_0(x) dx \\ &\leq \left(r + \frac{a}{2} \left((\theta - \bar{x}_0)^2 + \sigma_0^2 \right) + c N_0 \right) N_0, \end{aligned} \quad (58)$$

where the initial phenotypic mean \bar{x}_0 , phenotypic variance σ_0^2 and total abundance N_0 are finite by our assumptions on $\nu_0(x)$. To simplify analysis, we further assume

$$\nu_0(x) = \frac{N_0}{\sqrt{2\pi\sigma_0^2}} \left(-\frac{(x - \bar{x}_0)^2}{2\sigma_0^2} \right). \quad (59)$$

If some function $f(x, t)$ that maps $\mathbb{R} \times [0, \infty)$ to \mathbb{R} satisfies, for each fixed $t \geq 0$, $f(\cdot, t) \in C_1(\mathbb{R})^+$, then by $m(f, x)$ we mean $m(f(\cdot, t), x)$ for each $t \geq 0$. With this convention, consider the following initial value problem (IVP):

$$\begin{cases} \dot{\nu}(x, t) = m(\nu, x) \nu(x, t) + \frac{\mu}{2} \Delta \nu(x, t), & (x, t) \in \mathbb{R} \times (0, \infty) \\ \nu(x, t) = \nu_0(x), & (x, t) \in \mathbb{R} \times \{0\} \end{cases} \quad (60)$$

To solve for the lower bound of $\nu(x, t)$, let us assume $m(f, x) = r - \frac{a}{2}(\theta - x)^2 - c \int_{\mathbb{R}} f(x) dx$. Then

$$\bar{m}(t) = \frac{\int_{\mathbb{R}} m(\nu, x) \nu(x, t) dx}{\int_{\mathbb{R}} \nu(x, t) dx} = r - \frac{a}{2} \left((\theta - \bar{x}(t))^2 + \sigma^2(t) \right) - c N(t). \quad (61)$$

If the solution $\nu(x, t)$ to the IVP (60) is Gaussian, then we can derive differential equations for \bar{x} , σ^2 and N as in the main text to get

$$\dot{\bar{x}} = \sigma^2 \left(\frac{\partial \bar{m}}{\partial \bar{x}} - \frac{\partial \bar{m}}{\partial \bar{x}} \right) = a \sigma^2 (\theta - \bar{x}) \quad (62)$$

$$\dot{\sigma}^2 = \mu + 2\sigma^4 \left(\frac{\partial \bar{m}}{\partial \sigma^2} - \frac{\partial \bar{m}}{\partial \sigma^2} \right) = \mu - a \sigma^4 \quad (63)$$

$$\dot{N} = \bar{m} N = \left[r - \frac{a}{2} \left((\theta - \bar{x})^2 + \sigma^2 \right) - c N \right] N. \quad (64)$$

Under these assumptions, we can calculate the time derivative of $\nu(x, t)$ as

$$\begin{aligned}
\dot{\nu}(x, t) &= \left(\frac{\dot{N}}{N} + \frac{2\dot{x}(x - \bar{x})\sigma^2 + ((x - \bar{x})^2 - \sigma^2)\dot{\sigma}^2}{2\sigma^4} \right) \nu(x, t) \\
&= \left(\bar{m} + a(x - \bar{x})(\theta - \bar{x}) + \frac{\mu}{2} \frac{((x - \bar{x})^2 - \sigma^2)}{\sigma^4} - \frac{a}{2} ((x - \bar{x})^2 - \sigma^2) \right) \nu(x, t) \\
&= \left(r - \frac{a}{2} [(\theta - \bar{x})^2 + (x - \bar{x})^2] + a(x - \bar{x})(\theta - \bar{x}) - cN \right) \nu(x, t) + \frac{\mu}{2} \Delta \nu(x, t) \\
&= \left(r - \frac{a}{2} (\theta - x)^2 - cN \right) \nu(x, t) + \frac{\mu}{2} \Delta \nu(x, t) = m(\nu, x) \nu(x, t) + \frac{\mu}{2} \Delta \nu(x, t). \quad (65)
\end{aligned}$$

Thus, IVP (60) admits a Gaussian solution. To get the upper bound on $\nu(x, t)$, one can perform the same series of calculations to find that IVP (60) admits a Gaussian solution when $c = 0$. Hence, for each $t \geq 0$ $\nu(x, t)$ will have finite phenotypic mean, variance and total abundance for general m satisfying inequality (70).

5.2 The relation between diffusion and convolution with a Gaussian kernel

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

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

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

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

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

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

Hence, by the fundamental theorem of calculus,

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

In particular,

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

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

Picking up from the main text §2.1,

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

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

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

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

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

In the Gaussian case, we also have

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

5.4 Numerical evidence of finite moments

Here we use a numerical argument to suggest, for

$$r - \frac{a}{2}(\theta - x)^2 - c \int_{\mathbb{R}} f(x) dx \leq m(f, x) \leq r - \frac{a}{2}(\theta - x)^2, \quad \forall (f, x) \in C_1(\mathbb{R})^+ \times \mathbb{R}, \tag{70}$$

the density process $\nu(x, t)$ defined by SPDE (??) of the main text satisfies $\int_{\mathbb{R}} |x|^n \nu(x, t) dx < \infty$ for each $n \in \mathbb{N}$.

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

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

$$\mathbb{E} \left(\int_{\mathbb{R}} |h(x)| \nu(x, t) dx \right) < \infty, \quad (71)$$

$$\mathbb{E} \left(\int_{\mathbb{R}} h^2(x) \nu(x, t) dx \right) < \infty, \quad (72)$$

$$\mathbb{E} \left(\int_0^t \int_{\mathbb{R}} \nu(x, s) |h(x)| \left| m(\nu, x) f(x) + \frac{\mu}{2} \Delta f(x) \right| dx ds \right) < \infty, \quad \forall f \in C_b(\mathbb{R}). \quad (73)$$

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

$$\star := \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E}[H(t + \varepsilon) - H(t) | \nu(x, t) = \nu_1(x, t)] \quad (74)$$

$$= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E} \left[\int_{\mathbb{R}} h(x) (\nu(x, t + \varepsilon) - \nu(x, t)) dx \middle| \nu(x, t) = \nu_1(x, t) \right] \quad (75)$$

$$= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E} \left[\int_{\mathbb{R}} \int_t^{t+\varepsilon} f(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) + \sqrt{V \nu(x, s)} \dot{W}(x, s) \right) ds dx \middle| \nu(x, t) = \nu_1(x, t) \right]. \quad (76)$$

By assumption (73) we can use Fubini's theorem to write, with probability one,

$$\int_{\mathbb{R}} \int_t^{t+\varepsilon} h(x) (m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s)) ds dx = \int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) (m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s)) dx ds. \quad (77)$$

By assumption (72) we have $\varphi(x, t) = h(x) \sqrt{V \nu(x, t)}$ implies $\varphi \in L_C^2(\mathbb{R} \times [0, \infty))$. Is it true by definition that

$$\begin{aligned} & \mathbb{E} \left[\int_{\mathbb{R}} \int_t^{t+\varepsilon} h(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) ds dx \middle| \nu(x, t) = \nu_1(x, t) \right] \\ &= \mathbb{E} \left[\int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) dx ds \middle| \nu(x, t) = \nu_1(x, t) \right] = 0? \end{aligned} \quad (78)$$

Hence,

$$\star = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E} \left[\int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) \right) dx ds \middle| \nu(x, t) = \nu_1(x, t) \right]. \quad (79)$$

By assumption (73) we know that there exists a $\delta > 0$ such that for each positive $\varepsilon < \delta$ and with probability one

$$\begin{aligned} & \left| \int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) \right) dx ds \right| \\ & \leq \int_t^{t+\delta} \int_{\mathbb{R}} |h(x)| \left| m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) \right| dx ds < \infty \end{aligned} \quad (80)$$

almost surely. Thus, by Lebesgue's dominated convergence theorem, the drift component of the process $F(t)$ can be computed as

$$\begin{aligned} \star &= \mathbb{E} \left[\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) \right) dx ds \middle| \nu(x, t) = \nu_1(x, t) \right] \\ &= \int_{\mathbb{R}} h(x) \left(m(\nu_1, x) \nu_1(x, t) + \frac{\mu}{2} \Delta \nu_1(x, t) \right) dx. \end{aligned} \quad (81)$$

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

$$\begin{aligned} \star \star &:= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{V} [H(t + \varepsilon) - H(t) | \nu(x, t) = \nu_1(x, t)] \\ &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{V} \left[\int_{\mathbb{R}} h(x) (\nu(x, t + h) - \nu(x, t)) dx \middle| \nu(x, t) = \nu_1(x, t) \right]. \end{aligned} \quad (82)$$

We can rewrite the integral inside expression (82) as

$$\int_{\mathbb{R}} \int_t^{t+\varepsilon} h(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) + \sqrt{V \nu(x, s)} \dot{W}(x, s) \right) ds dx. \quad (83)$$

By our assumption (72) that $\int_{\mathbb{R}} h^2(x) \nu(x, t) dx < \infty$ for each $t \geq 0$, we have $\varphi(x, t) = h(x) \sqrt{V \nu(x, t)}$ implies $\varphi \in L_C^2(\mathbb{R} \times [0, \infty))$. We have already found

$$\begin{aligned} \mathbb{E} \left[\int_{\mathbb{R}} \int_t^{t+\varepsilon} f(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) + \sqrt{V \nu(x, s)} \dot{W}(x, s) \right) ds dx \middle| \nu(x, t) = \nu_1(x, t) \right] \\ = \mathbb{E} \left[\int_t^{t+\varepsilon} \int_{\mathbb{R}} f(x) \left(m(\nu, x) \nu(x, s) + \frac{\mu}{2} \Delta \nu(x, s) \right) dx ds \middle| \nu(x, t) = \nu_1(x, t) \right]. \end{aligned} \quad (84)$$

Then, since $h^2(x) \nu(x, t)$ is integrable by assumption (72), we have

$$\begin{aligned} \star \star &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E} \left[\left(\int_t^{t+\varepsilon} \int_{\mathbb{R}} h(x) \sqrt{V \nu(x, s)} \dot{W}(x, s) ds dx \right)^2 \middle| \nu(x, t) = \nu_1(x, t) \right] \\ &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E} \left[\int_t^{t+\varepsilon} \int_{\mathbb{R}} V h^2(x) \nu(x, s) dx ds \middle| \nu(x, t) = \nu_1(x, t) \right]. \end{aligned} \quad (85)$$

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

$$\int_t^{t+\varepsilon} \int_{\mathbb{R}} V h^2(x) \nu(x, s) dx ds \leq \int_t^{t+\delta} \int_{\mathbb{R}} V h^2(x) \nu(x, s) dx ds. \quad (86)$$

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

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

Thus, using standard SDE notation, we have

$$dH(t) = \left(\int_{\mathbb{R}} f(x) \left(m(\nu, x) \nu(x, t) + \frac{\mu}{2} \Delta \nu(x, t) \right) dx \right) dt + \sqrt{V \int_{\mathbb{R}} h(x) \nu(x, t) dx} dW_H(t) \quad (88)$$

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

5.5.1 Derivation for trait mean

Setting $\tilde{x}(t) = \int_{\mathbb{R}} x \nu(x, t)$ and applying formula (88) provides

$$d\tilde{x} = \left(\bar{x} m N + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta \nu dx \right) dt + \sqrt{V \bar{x}^2 N} d\tilde{W}_2, \quad (89)$$

where,

$$d\tilde{W}_2 = d\hat{\mathbf{W}}_{\sqrt{V \bar{x}^2 N}} = \frac{1}{\sqrt{V \bar{x}^2 N}} \int_{\mathbb{R}} x \sqrt{V \nu(x, t)} \dot{W}(x, t) dx dt. \quad (90)$$

Using Itô's (finite-dimensional) quotient rule on $\bar{x} = \tilde{x}/N$, we obtain

$$\begin{aligned} d\bar{x} &= d \left(\frac{\tilde{x}}{N} \right) = \frac{\tilde{x}}{N} \left(\frac{d\tilde{x}}{\tilde{x}} - \frac{dN}{N} - \frac{d\tilde{x}}{\tilde{x}} \frac{dN}{N} + \left(\frac{dN}{N} \right)^2 \right) \\ &= \frac{d\tilde{x}}{N} - \bar{x} \frac{dN}{N} - \frac{d\tilde{x}}{N} \frac{dN}{N} + \bar{x} \left(\frac{dN}{N} \right)^2 \\ &= \bar{x} m dt + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x, t) dx + \frac{V}{2} \bar{x}^2 + \sqrt{V \frac{\bar{x}^2}{N}} d\tilde{W}_2 - \bar{x} m dt - V \sqrt{\frac{\bar{x}}{N}} dW_1 - V \frac{\bar{x}}{N} dt + V \frac{\bar{x}}{N} dt \\ &= \left(\text{Cov}_t(x, m) + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x, t) dx \right) dt + \sqrt{V \frac{\bar{x}^2}{N}} d\tilde{W}_2 - \bar{x} \sqrt{\frac{V}{N}} dW_1. \end{aligned} \quad (91)$$

Note that

$$\begin{aligned} \sqrt{V \frac{\bar{x}^2}{N}} d\tilde{W}_2 - \bar{x} \sqrt{\frac{V}{N}} dW_1 &= \frac{1}{N} \int_{\mathbb{R}} x \sqrt{V \nu(x, t)} \dot{W}(x, t) dx - \frac{\bar{x}}{N} \int_{\mathbb{R}} \sqrt{V \nu(x, t)} \dot{W}(x, t) dx \\ &= \int_{\mathbb{R}} \frac{x - \bar{x}}{N} \sqrt{V \nu(x, t)} \dot{W}(x, t) dx \end{aligned} \quad (92)$$

and

$$\mathbb{V} \left(\int_{\mathbb{R}} \frac{x - \bar{x}}{N} \sqrt{V \nu(x, t)} \dot{W}(x, t) dx \right) = \frac{V}{N} \int_{\mathbb{R}} (x - \bar{x})^2 p(x, t) dx = V \frac{\sigma^2}{N}. \quad (93)$$

Hence, by setting

$$dW_2 = \frac{\int_{\mathbb{R}} \frac{(x-\bar{x})}{N} \sqrt{V\nu(x,t)} \dot{W}(x,t) dx}{\sqrt{V\sigma^2/N}} \quad (94)$$

we can write

$$d\bar{x} = \left(\text{Cov}_t(x, m) + \frac{\mu}{2} \int_{\mathbb{R}} x \Delta p(x, t) dx \right) dt + \sqrt{V \frac{\sigma^2}{N}} dW_2. \quad (95)$$

5.5.2 Derivation for trait variance

UNDER CONSTRUCTION

Setting $\tilde{\sigma}^2(t) = \int_{\mathbb{R}} x^2 \nu(x, t) dx$ and applying formula (88) provides

$$d\tilde{\sigma}^2 = \left(\overline{x^2 m} N + \frac{\mu}{2} \int_{\mathbb{R}} x^2 \Delta \nu dx \right) dt + \sqrt{V \overline{x^4} N} d\tilde{W}_3 \quad (96)$$

where

$$d\tilde{W}_3 = d\hat{\mathbf{W}}_{\sqrt{V \overline{x^4} N}} = \frac{1}{\sqrt{V \overline{x^4} N}} \int_{\mathbb{R}} x^2 \sqrt{V \nu(x, t)} \dot{W}(x, t) dx. \quad (97)$$

Use Itô's quotient rule on $\tilde{\sigma}^2 = \tilde{\sigma}^2/N$ to obtain:

$$d\tilde{\sigma}^2 = d\left(\frac{\tilde{\sigma}^2}{N}\right) = \frac{\tilde{\sigma}^2}{N} \left(\frac{d\tilde{\sigma}^2}{\tilde{\sigma}^2} - \frac{dN}{N} - \frac{d\tilde{\sigma}^2}{\tilde{\sigma}^2} \frac{dN}{N} + \left(\frac{dN}{N}\right)^2 \right) \quad (98)$$

$$= \frac{d\tilde{\sigma}^2}{N} - \tilde{\sigma}^2 \frac{dN}{N} - \frac{d\tilde{\sigma}^2}{N} \frac{dN}{N} + \tilde{\sigma}^2 \left(\frac{dN}{N}\right)^2. \quad (99)$$

Table 1 of the main text implies

$$\left(\sqrt{V \frac{\overline{x^4}}{N}} d\tilde{W}_3 \right) \left(\sqrt{\frac{V}{N}} dW_1 \right) = \frac{\int_{\mathbb{R}} \sqrt{x^4 V \nu(x, t)} \sqrt{V \nu(x, t)} dx}{N^2} dt = V \frac{\overline{x^2}}{N} dt. \quad (100)$$

and hence

$$d\tilde{\sigma}^2 = \left(\overline{x^2 m} + \frac{\mu}{2} \int_{\mathbb{R}} x^2 \Delta p(x, t) dx \right) dt + \sqrt{V \frac{\overline{x^4}}{N}} d\tilde{W}_3 - \tilde{\sigma}^2 \bar{m} dt + \tilde{\sigma}^2 \sqrt{\frac{V}{N}} dW_1 - V \frac{\overline{x^2}}{N} dt - \tilde{\sigma}^2 \frac{V}{N} dt. \quad (101)$$

$$= \left(\text{Cov}_t(x^2, m) + \frac{\mu}{2} \int_{\mathbb{R}} x^2 \Delta p dx \right) dt + \sqrt{V \frac{\overline{x^4}}{N}} d\tilde{W}_3 - \tilde{\sigma}^2 \sqrt{\frac{V}{N}} dW_1. \quad (102)$$

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

$$d\sigma^2 = d\tilde{\sigma}^2 - 2\bar{x}d\bar{x} - (d\bar{x})^2 \quad (103)$$

$$\begin{aligned}
&= \left(\text{Cov}_t(x^2, m) + \mu \right) dt + \sqrt{\frac{Vx^4}{N}} d\tilde{W}_3 - \tilde{\sigma}^2 \sqrt{\frac{V}{N}} dW_1 \\
&\quad - 2\bar{x} \left(\text{Cov}_t(x, m) dt + \sqrt{\frac{V\sigma^2}{N}} dW_2 \right) - \left(\text{Cov}_t(x, m) dt + \sqrt{\frac{V\sigma^2}{N}} dW_2 \right)^2 \\
&= \left(\text{Cov}_t(x^2 - 2\bar{x}x, m) + \mu \right) dt + \sqrt{\frac{Vx^4}{N}} d\tilde{W}_3 - \tilde{\sigma}^2 \sqrt{\frac{V}{N}} dW_1 - 2\bar{x} \sqrt{\frac{V\sigma^2}{N}} dW_2 - \left(\frac{V\sigma^2}{N} \right) dt \\
&= \left(\text{Cov}_t(x - \bar{x})^2, m \right) + \mu - \frac{V\sigma^2}{N} \right) dt + \sqrt{\frac{Vx^4}{N}} d\tilde{W}_3 - \tilde{\sigma}^2 \sqrt{\frac{V}{N}} dW_1 - 2\bar{x} dW_2
\end{aligned}$$

649 In light of

$$\begin{aligned}
\sqrt{\frac{Vx^4}{N}} d\tilde{W}_3 - \tilde{\sigma}^2 \sqrt{\frac{V}{N}} dW_1 - 2\bar{x} dW_2 &= \frac{1}{N} \int_{\mathbb{R}} (x^2 - \tilde{\sigma}^2 - 2\bar{x}(x - \bar{x})) \sqrt{V\nu(x, t)} \dot{W}(x, t) dx \\
&= \frac{1}{N} \int_{\mathbb{R}} ((x - \bar{x})^2 - \sigma^2) \sqrt{V\nu(x, t)} \dot{W}(x, t) dx
\end{aligned} \tag{104}$$

650 we set

$$\mathcal{V}^2 = \mathbb{V} \left(\frac{1}{N} \int_{\mathbb{R}} ((x - \bar{x})^2 - \sigma^2) \sqrt{V\nu(x, s)} \dot{W}(x, s) dx \right) \tag{105}$$

$$= \frac{V}{N} \left(\int_{\mathbb{R}} ((x - \bar{x})^4 - 2(x - \bar{x})^2 \sigma^2 + \sigma^4) p(x, t) dx \right) \tag{106}$$

$$= \frac{V}{N} \left(\overline{(x - \bar{x})^4} - \sigma^4 \right) \tag{107}$$

651 and

$$dW_3 = \frac{1}{N\mathcal{V}} \int_{\mathbb{R}} ((x - \bar{x})^2 - \sigma^2) \sqrt{V\nu(x, t)} \dot{W}(x, t) dx \tag{108}$$

652 so that

$$d\sigma^2 = \left(\text{Cov}_t((x - \bar{x})^2, m) + \frac{\mu}{2} \int_{\mathbb{R}} (x - \bar{x})^2 \Delta \nu(x, t) dx - V \frac{\sigma^2}{N} \right) dt + \sqrt{V \frac{(x - \bar{x})^4 - \sigma^4}{N}} dW_3. \tag{109}$$

653 Table 1 of the main text implies

$$dW_1 dW_2 = \frac{\int_{\mathbb{R}} (x - \bar{x}) \nu(x, t) dx}{\sqrt{N\sigma^2}} = 0, \tag{110}$$

$$dW_1 dW_3 = \frac{\int_{\mathbb{R}} ((x - \bar{x})^2 - \sigma^2) \nu(x, t) dx}{\sqrt{(x - \bar{x})^4 - \sigma^4}} = 0, \tag{111}$$

$$dW_2 dW_3 = \frac{\int_{\mathbb{R}} (x - \bar{x}) ((x - \bar{x})^2 - \sigma^2) \nu(x, t) dx}{\sqrt{\sigma^2 ((x - \bar{x})^4 - \sigma^4)}} = \frac{N(\overline{x - \bar{x}})^3}{\sqrt{\sigma^2 ((x - \bar{x})^4 - \sigma^4)}}. \quad (112)$$

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

5.6 Derivation of diffuse coevolution model

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

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

$$\mathcal{B}_{ij}(X) = C_i^{\sum_{l \neq j} \mathcal{O}_{ii}(x_{ij}, x_{il}) + \sum_{k \neq i} \sum_{l=1}^{n_k} \mathcal{O}_{ik}(x_{ij}, x_{kl})}, \quad (113)$$

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

$$\mathcal{W}_{ij}(X) = \mathcal{A}_i(x_{ij}) \mathcal{B}_{ij}(X), \quad (114)$$

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. The Malthusian fitness m is related to \mathcal{W} by $m = \ln \mathcal{W}$ (Crow and Kimura 2009). Hence,

$$\begin{aligned} m_{ij}(X) &= \ln \mathcal{A}_i(x_{ij}) + \ln \mathcal{B}_{ij}(X) \\ &= \ln \mathcal{A}_i(x_{ij}) + \left(\sum_{l \neq j} \mathcal{O}_{ii}(x_{ij}, x_{il}) + \sum_{k \neq i} \sum_{l=1}^{n_k} \mathcal{O}_{ik}(x_{ij}, x_{kl}) \right) \ln C_i. \end{aligned} \quad (115)$$

In the diffusion limit we rescale individual mass of species i by $1/n_i$. We set $\mathbf{n} = (n_1, \dots, n_S)$ and write the rescaled limit of the Malthusian fitness as

$$m_{ij}^{(\mathbf{n})}(X_t) = \ln \mathcal{A}_i(x_{ij}) + \left(\frac{1}{n_i} \sum_{l \neq j} \mathcal{O}_{ii}(x_{ij}, x_{il}) + \sum_{k \neq i} \frac{1}{n_k} \sum_{l=1}^{n_k} \mathcal{O}_{ik}(x_{ij}, x_{kl}) \right) \ln C_i. \quad (116)$$

Denoting the density process describing the distribution of trait values in species i by $\nu_i(x, t)$ and writing $\boldsymbol{\nu} = (\nu_1, \dots, \nu_S)$, the diffusion limit of Malthusian fitness for trait value x in species i at time $t \geq 0$ becomes

$$\begin{aligned}
m_i(\boldsymbol{\nu}, x) &= \ln \mathcal{A}_i(x_{ij}) \\
&+ \left(\int_{\mathbb{R}} \mathcal{O}_{ii}(x_{ij}, x) \nu_i(x, t) dx - \frac{\int_{\mathbb{R}} \mathcal{O}_{ii}(x_{ij}, x) \nu_i(x, t) dx}{\int_{\mathbb{R}} \nu_i(x, t) dx} + \sum_{k \neq i} \int_{\mathbb{R}} \mathcal{O}_{ik}(x_{ij}, x) \nu_k(x, t) dx \right) \ln C_i \\
&= \ln \mathcal{A}_i(x_{ij}) + \left(-\bar{\mathcal{O}}_{ii}(x, t_{ij}) + \sum_{k=1}^S N_k(t) \bar{\mathcal{O}}_{ik}(x, t_{ij}) \right) \ln C_i. \quad (117)
\end{aligned}$$

When $\nu_i(x, t)$ and $\nu_j(x, t)$ are Gaussian curves, the expected niche overlap between an individual at niche location x with individuals in species j is

$$\bar{\mathcal{O}}_{ij}(x, t) = \frac{\int_{\mathbb{R}} \mathcal{O}_{ij}(x, x_j) \nu_j(x_j, t) dx_j}{\int_{\mathbb{R}} \nu_j(x_j, t) dx_j} = \frac{1}{\sqrt{2\pi(w_i + w_j + \sigma_j^2(t))}} \exp\left(-\frac{(x - \bar{x}_j(t))^2}{2(w_i + w_j + \sigma_j^2(t))}\right). \quad (118)$$

Setting

$$r_i = \ln\left(\frac{Q_i U_i}{\sqrt{1 + A_i w_i}}\right), \quad (119a)$$

$$a_i = \frac{A_i}{1 + A_i w_i}, \quad (119b)$$

$$\tilde{b}_{ij}(t) = \frac{1}{w_i + w_j + \sigma_j^2(t)}, \quad (119c)$$

$$c_i = -\ln C_i, \quad (119d)$$

we get

$$m_i(\boldsymbol{\nu}, x) = r_i - \frac{a_i}{2}(x - \theta_i)^2 + c_i \left(\sqrt{\frac{\tilde{b}_{ii}(t)}{2\pi}} e^{-\frac{\tilde{b}_{ii}(t)}{2}(x - \bar{x}_i(t))^2} - \sum_{j=1}^S N_j(t) \sqrt{\frac{\tilde{b}_{ij}(t)}{2\pi}} e^{-\frac{\tilde{b}_{ij}(t)}{2}(x - \bar{x}_j(t))^2} \right). \quad (120)$$

For the remainder of the derivation we suppress notation indicating dependency on t and x . From (120) we calculate

$$\frac{\partial m_i}{\partial \bar{x}_i} = c_i \tilde{b}_{ii}(x - \bar{x}_i) \sqrt{\frac{\tilde{b}_{ii}}{2\pi}} e^{-\frac{\tilde{b}_{ii}}{2}(x - \bar{x}_i)^2} \quad (121)$$

$$\frac{\partial m_i}{\partial G_i} = \frac{c_i ((x - \bar{x}_i)^2 - \sigma_i^2 - 2w_i)}{2(\sigma_i^2 + 2w_i)^2} \sqrt{\frac{\tilde{b}_{ii}}{2\pi}} e^{-\frac{\tilde{b}_{ii}}{2}(x - \bar{x}_i)^2} \quad (122)$$

and hence

$$\frac{\partial m_i}{\partial \bar{x}_i} = 0, \quad (123)$$

$$\begin{aligned}\frac{\partial \bar{m}_i}{\partial G_i} &= \frac{c_i}{2(\sigma_i^2 + 2w_i)^2} \left(\frac{(\sigma_i^2 + 2w_i)\sigma_i^2}{2(w_i + \sigma_i^2)} - \sigma_i^2 - 2w_i \right) \sqrt{\frac{b_{ii}}{2\pi}} \\ &= -\frac{c_i}{2(\sigma_i^2 + 2w_i)^2} \frac{(\sigma_i^2 + 2w_i)^2}{2(\sigma_i^2 + w_i)} \sqrt{\frac{b_{ii}}{2\pi}} = -\frac{c_i b_{ii}}{2} \sqrt{\frac{b_{ii}}{2\pi}},\end{aligned}\quad (124)$$

where

$$b_{ij} = \frac{1}{w_i + w_j + \sigma_i^2 + \sigma_j^2}.\quad (125)$$

The average fitness for species i is

$$\bar{m}_i = r_i - \frac{a_i}{2} \left((\bar{x}_i - \theta_i)^2 + G_i + \eta_i N_i \right) + c_i \left(\sqrt{\frac{b_{ii}}{2\pi}} - \sum_{j=1}^S N_j \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2} \right).\quad (126)$$

Hence

$$\frac{\partial \bar{m}_i}{\partial \bar{x}_i} = a_i(\theta_i - \bar{x}_i) - c_i \sum_j N_j b_{ij}(\bar{x}_j - \bar{x}_i) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2},\quad (127)$$

$$\frac{\partial \bar{m}_i}{\partial G_i} = -\frac{a_i}{2} - \frac{c_i}{2} \left(b_{ii} \sqrt{\frac{b_{ii}}{2\pi}} - \sum_{j=1}^S N_j b_{ij} (1 - b_{ij}(\bar{x}_i - \bar{x}_j)^2) \sqrt{\frac{b_{ij}}{2\pi}} e^{-\frac{b_{ij}}{2}(\bar{x}_i - \bar{x}_j)^2} \right).\quad (128)$$

Applying equations (26), (40) and (41) recovers system (46) of the main text.

5.7 Moments of competition coefficients and selection gradients

The calculations are straightforward applications of Gaussian products. That is, if

$$f_1(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{(\mu_1 - x)^2}{2\sigma_1^2}\right), \quad f_2(x) = \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{(\mu_2 - x)^2}{2\sigma_2^2}\right),\quad (129)$$

then

$$f_1(x)f_2(x) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left(-\frac{(\mu_1 - \mu_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right) \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}} \exp\left(-\frac{(\tilde{\mu} - x)^2}{2\tilde{\sigma}^2}\right),\quad (130)$$

where

$$\tilde{\mu} = \frac{\sigma_2^2 \mu_1 + \sigma_1^2 \mu_2}{\sigma_1^2 + \sigma_2^2}, \quad \tilde{\sigma}^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}.\quad (131)$$

Hence:

$$\begin{aligned}
\bar{\alpha} &= \int_{\mathbb{R}} \int_{\mathbb{R}} \alpha(\bar{x}_i, \bar{x}_j) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \frac{c}{\bar{r}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi(b^{-1} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(b^{-1} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j = \frac{c/\bar{r}}{\sqrt{2\pi(b^{-1} + 2V_{\bar{X}})}}, \quad (132)
\end{aligned}$$

$$\begin{aligned}
\bar{\beta} &= \int_{\mathbb{R}} \int_{\mathbb{R}} \beta(\bar{x}_i, \bar{x}_j) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= cNb \int_{\mathbb{R}} (\bar{x} - \bar{x}_j) \frac{1}{\sqrt{2\pi(b^{-1} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(b^{-1} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j = 0, \quad (133)
\end{aligned}$$

$$\begin{aligned}
\bar{\gamma} &= \int_{\mathbb{R}} \int_{\mathbb{R}} \gamma(\bar{x}_i, \bar{x}_j) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= cNb \int_{\mathbb{R}} \left(1 - b((\bar{x} - \bar{x}_j)^2 + V_{\bar{X}})\right) \frac{1}{\sqrt{2\pi(b^{-1} + V_{\bar{X}})}} \exp\left(-\frac{(\bar{x} - \bar{x}_j)^2}{2(b^{-1} + V_{\bar{X}})}\right) f_{\bar{X}}(\bar{x}_j) d\bar{x}_j \\
&= \frac{cNb(1 - 2bV_{\bar{X}})}{\sqrt{2\pi(b^{-1} + 2V_{\bar{X}})}} \quad (134)
\end{aligned}$$

$$\begin{aligned}
\text{Var}_{f_{\bar{X}}}(\alpha) &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{\alpha} - \alpha(\bar{x}_i, \bar{x}_j))^2 f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \frac{c^2}{\bar{r}^2} \int_{\mathbb{R}} \int_{\mathbb{R}} \left(\frac{1}{\sqrt{2\pi(b^{-1} + 2V_{\bar{X}})}} - \sqrt{\frac{b}{2\pi}} \exp\left(-\frac{b}{2}(\bar{x}_i - \bar{x}_j)^2\right) \right)^2 f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \frac{c^2}{\bar{r}^2} \left(\frac{1}{2\pi(b^{-1} + 2V_{\bar{X}})} - \frac{1}{\pi(b^{-1} + 2V_{\bar{X}})} + \sqrt{\frac{b}{2\pi}} \frac{1}{\sqrt{2\pi(b^{-1} + 2V_{\bar{X}})}} \right) \\
&= \frac{c^2}{2\pi\bar{r}^2} \left(\frac{1 + 2bV_{\bar{X}}}{b^{-1} + 2V_{\bar{X}}} - \frac{1}{b^{-1} + 2V_{\bar{X}}} \right) = \frac{c^2}{2\pi\bar{r}^2} \frac{2bV_{\bar{X}}}{b^{-1} + 2V_{\bar{X}}}, \quad (135)
\end{aligned}$$

$$\begin{aligned}
\text{Var}_{f_{\bar{X}}}(\beta) &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{\beta} - \beta(\bar{x}_i, \bar{x}_j))^2 f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} c^2 N^2 b^2 (\bar{x}_i - \bar{x}_j)^2 \frac{b}{2\pi} \exp\left(-\frac{b}{2}(\bar{x}_i - \bar{x}_j)^2\right) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \sqrt{\frac{b}{2\pi}} c^2 N^2 b^2 (b^{-1} + 2V_{\bar{X}}), \quad (136)
\end{aligned}$$

$$\begin{aligned}
\text{Var}_{f_{\bar{X}}}(\gamma) &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{\gamma} - \gamma(\bar{x}_i, \bar{x}_j))^2 f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \quad (137)
\end{aligned}$$

$$\begin{aligned}
\text{Cov}_{f_{\bar{X}}}(\alpha, \beta) &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{\alpha} - \alpha(\bar{x}_i, \bar{x}_j)) (\bar{\beta} - \beta(\bar{x}_i, \bar{x}_j)) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\
&= \quad (138)
\end{aligned}$$

$$\begin{aligned} \text{Cov}_{f_{\bar{X}}}(\alpha, \gamma) &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{\alpha} - \alpha(\bar{x}_i, \bar{x}_j)) (\bar{\gamma} - \gamma(\bar{x}_i, \bar{x}_j)) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\ &= \end{aligned} \quad (139)$$

$$\begin{aligned} \text{Cov}_{f_{\bar{X}}}(\gamma, \beta) &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\bar{\gamma} - \gamma(\bar{x}_i, \bar{x}_j)) (\bar{\beta} - \beta(\bar{x}_i, \bar{x}_j)) f_{\bar{X}}(\bar{x}_i) f_{\bar{X}}(\bar{x}_j) d\bar{x}_i d\bar{x}_j \\ &= \end{aligned} \quad (140)$$

References

- Abdala-Roberts, Luis, and Kailen A. Mooney. 2014. “Ecological and Evolutionary Consequences of Plant Genotype Diversity in a Tri-Trophic System.” *Ecology* 95 (10). Wiley: 2879–93.
- Abrams, Peter. 1980. “Some Comments on Measuring Niche Overlap.” *Ecology* 61 (1). Wiley: 44–49.
- Ågren, Göran I., and Folke O. Andersson. 2012. *Terrestrial Ecosystem Ecology: Principles and Applications*. Cambridge University Press.
- Barton, N.H., A.M. Etheridge, and A. Véber. 2017. “The Infinitesimal Model: Definition, Derivation, and Implications.” *Theoretical Population Biology* 118 (December). Elsevier BV: 50–73.
- Barton, Nick, and Alison Etheridge. 2019. “Mathematical Models in Population Genetics.” Wiley.
- Barton, Nick, Alison Etheridge, and Amandine Véber. 2013. “Modelling Evolution in a Spatial Continuum.” *Journal of Statistical Mechanics: Theory and Experiment* 2013 (01). IOP Publishing: P01002.
- Bertoin, Jean, and Jean-François Le Gall. 2003. “Stochastic Flows Associated to Coalescent Processes.” *Probability Theory and Related Fields* 126 (2). Springer Science; Business Media: 261–88.
- Bulmer. 1980. *The Mathematical Theory of Quantitative Genetics*. Oxford University Press.
- Bürger, Reinhard. 2000. *The Mathematical Theory of Selection, Recombination, and Mutation*. Wiley.
- Chesson, Peter. 2000. “Mechanisms of Maintenance of Species Diversity.” *Annual Review of Ecology and Systematics* 31 (1). Annual Reviews: 343–66.
- Conner, Jeffrey K. 2004. *A Primer of Ecological Genetics*. Sinauer Associates.
- Crow, James F., and Motoo Kimura. 2009. *An Introduction to Population Genetics Theory*. The Blackburn Press.
- Crutsinger, Gregory M. 2015. “A Community Genetics Perspective: Opportunities for the Coming Decade.” *New Phytologist* 210 (1). Wiley: 65–70.
- David, Thomas I., Jonathan Storkey, and Carly J. Stevens. 2019. “Understanding How Changing Soil Nitrogen Affects Plantpollinator Interactions.” *Arthropod-Plant Interactions* 13 (5). Springer Science; Business Media LLC: 671–84.
- Dawson, Donald A. 1975. “Stochastic Evolution Equations and Related Measure Processes.” *Journal of Multivariate Analysis* 5 (1). Elsevier BV: 1–52.
- Débarre, Florence, Sam Yeaman, and Frédéric Guillaume. 2015. “Evolution of Quantitative Traits Under a Migration-Selection Balance: When Does Skew Matter?” *The American Naturalist* 186 Suppl 1: S37–47.
- Etheridge, Alison M. 2000. *An Introduction to Superprocesses*. American Mathematical Society.
- . 2008. “Drift, Draft and Structure: Some Mathematical Models of Evolution.” In *Stochastic Models in Biological Sciences*. Institute of Mathematics Polish Academy of Sciences.
- Etheridge, Alison, and Peter March. 1991. “A Note on Superprocesses.” *Probability Theory and Related Fields* 89 (2). Springer: 141–47.
- Evans, Lawrence C. 2010. *Partial Differential Equations: Second Edition*. American Mathematical Society.
- . 2014. *An Introduction to Stochastic Differential Equations*. American Mathematical Society.
- Ewens, Warren J. 2004. *Mathematical Population Genetics*. Springer New York.
- 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.
- 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.

- 736 Fisher, R. A. 1923. "XXI.-on the Dominance Ratio." *Proceedings of the Royal Society of Edinburgh* 42.
737 Cambridge University Press: 321–41.
- 738 Fitzpatrick, Connor R., Anurag A. Agrawal, Nathan Basiliko, Amy P. Hastings, Marney E. Isaac, Michael
739 Preston, and Marc T. J. Johnson. 2015. "The Importance of Plant Genotype and Contemporary Evolution
740 for Terrestrial Ecosystem Processes." *Ecology* 96 (10). Wiley: 2632–42.
- 741 Fitzpatrick, Connor R., Anna V. Mikhailitchenko, Daniel N. Anstett, and Marc T. J. Johnson. 2017. "The
742 Influence of Range-Wide Plant Genetic Variation on Soil Invertebrate Communities." *Ecography* 41 (7).
743 Wiley: 1135–46.
- 744 Frank, S. A. 2012. "Natural Selection. IV. The Price Equation." *Journal of Evolutionary Biology* 25 (6).
745 Wiley: 1002–19.
- 746 Fridley, Jason D. 2017. "Plant Energetics and the Synthesis of Population and Ecosystem Ecology." Edited
747 by David Gibson. *Journal of Ecology* 105 (1). Wiley: 95–110.
- 748 Fussman, G. F., M. Loreau, and P. A. Abrams. 2007. "Eco-Evolutionary Dynamics of Communities and
749 Ecosystems." *Functional Ecology* 21 (3). Wiley: 465–77.
- 750 Harmon, Luke J., Cecilia S. Andreazzi, Florence Débarre, Jonathan Drury, Emma E. Goldberg, Ayana B.
751 Martins, Carlos J. Melián, et al. 2019. "Detecting the Macroevolutionary Signal of Species Interactions."
752 *Journal of Evolutionary Biology* 32 (8). Wiley: 769–82.
- 753 Harte, John. 2011. *Maximum Entropy and Ecology*. Oxford University Press.
- 754 Harte, John, and Erica A. Newman. 2014. "Maximum Information Entropy: A Foundation for Ecological
755 Theory." *Trends in Ecology & Evolution* 29 (7). Elsevier BV: 384–89.
- 756 Hickerson, M.J., B.C. Carstens, J. Cavender-Bares, K.A. Crandall, C.H. Graham, J.B. Johnson, L. Rissler,
757 P.F. Victoriano, and A.D. Yoder. 2010. "Phylogeography's Past, Present, and Future: 10 Years After Avise,
758 2000." *Molecular Phylogenetics and Evolution* 54 (1). Elsevier BV: 291–301.
- 759 Holt, Robert D. 1987. "On the Relation Between Niche Overlap and Competition: The Effect of Incommen-
760 surable Niche Dimensions." *Oikos* 48 (1). JSTOR: 110.
- 761 Kendall, David G. 1966. "Branching Processes Since 1873." *Journal of the London Mathematical Society*
762 s1-41 (1). Wiley: 385–406.
- 763 Kimmel, Marek, and David E. Axelrod. 2015. *Branching Processes in Biology*. Springer New York.
- 764 Kimura, M. 1965. "A Stochastic Model Concerning the Maintenance of Genetic Variability in Quantitative
765 Characters." *Proceedings of the National Academy of Sciences* 54 (3). Proceedings of the National Academy
766 of Sciences: 731–36.
- 767 Kimura, M., and J. F. Crow. 1978. "Effect of Overall Phenotypic Selection on Genetic Change at Individual
768 Loci." *Proceedings of the National Academy of Sciences* 75 (12). Proceedings of the National Academy of
769 Sciences: 6168–71.
- 770 Kirkpatrick, Mark, Toby Johnson, and Nick Barton. 2002. "General Models of Multilocus Evolution."
771 *Genetics* 161 (4). Genetics: 1727–50.
- 772 Konno, N., and T. Shiga. 1988. "Stochastic Partial Differential Equations for Some Measure-Valued Diffu-
773 sions." *Probability Theory and Related Fields* 79 (2). Springer Nature: 201–25.
- 774 Kopp, Michael, and Sergey Gavrillets. 2006. "Multilocus Genetics and the Coevolution of Quantitative Traits."
775 *Evolution* 60 (7). Wiley: 1321–36.
- 776 Kölzsch, Andrea, Adriana Alzate, Frederic Bartumeus, Monique de Jager, Ellen J. Weerman, Geerten M.
777 Hengeveld, Marc Naguib, Bart A. Nolet, and Johan van de Koppel. 2015. "Experimental Evidence for
778 Inherent Lévy Search Behaviour in Foraging Animals." *Proceedings of the Royal Society B: Biological Sciences*
779 282 (1807). The Royal Society: 20150424.

- 780 Kraft, Nathan J. B., William K. Cornwell, Campbell O. Webb, and David D. Ackerly. 2007. "Trait Evolution,
781 Community Assembly, and the Phylogenetic Structure of Ecological Communities." *The American Naturalist*
782 170 (2). University of Chicago Press: 271–83.
- 783 Lande, Russel. 1976. "Natural Selection and Random Genetic Drift in Phenotypic Evolution." *Evolution* 30
784 (2). Wiley: 314–34.
- 785 ———. 1980. "The Genetic Covariance between Characters Maintained by Pleiotropic Mutations." *Genetics*
786 94 (1): 203–15.
- 787 Lande, Russell, and Stevan J. Arnold. 1983. "The Measurement of Selection on Correlated Characters." *Evolution* 37 (6). JSTOR: 1210.
- 789 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.
- 791 Landis, Michael J., and Joshua G. Schraiber. 2017. "Pulsed Evolution Shaped Modern Vertebrate Body
792 Sizes." *Proceedings of the National Academy of Sciences* 114 (50). Proceedings of the National Academy of
793 Sciences: 13224–9.
- 794 Levins, Richard. 1968. *Evolution in Changing Environments: Some Theoretical Explorations. (MPB-2)*
795 *(Monographs in Population Biology)*. Princeton University Press.
- 796 Li, Zeng-Hu. 1998. "Absolute Continuity of Measure Branching Processes with Interaction." *Chinese Journal*
797 *of Applied Probability and Statistics* 14. Citeseer: 231–42.
- 798 Lion, Sébastien. 2018. "Theoretical Approaches in Evolutionary Ecology: Environmental Feedback as a
799 Unifying Perspective." *The American Naturalist* 191 (1). University of Chicago Press: 21–44.
- 800 Loreau, Michel. 2010. *From Populations to Ecosystems: Theoretical Foundations for a New Ecological*
801 *Synthesis*. Princeton University Press.
- 802 Lynch, Michael, and Bruce Walsh. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer Associates
803 is an imprint of Oxford University Press.
- 804 MacArthur, Robert H. 1969. "Species Packing, and what Competition Minimizes." *Proceedings of the*
805 *National Academy of Sciences* 64 (4). Proceedings of the National Academy of Sciences: 1369–71.
- 806 ———. 1970. "Species Packing and Competitive Equilibrium for Many Species." *Theoretical Population*
807 *Biology* 1 (1). Elsevier BV: 1–11.
- 808 ———. 1972. *Geographical Ecology*. Princeton University Press.
- 809 MacArthur, Robert H., and Richard Levins. 1967. "The Limiting Similarity, Convergence, and Divergence
810 of Coexisting Species." *The American Naturalist* 101 (921). University of Chicago Press: 377–85.
- 811 Manceau, Marc, Amaury Lambert, and Hélène Morlon. 2016. "A Unifying Comparative Phylogenetic
812 Framework Including Traits Coevolving Across Interacting Lineages." *Systematic Biology*, December. Oxford
813 University Press (OUP), syw115.
- 814 Marx, Hannah E., Cédric Dentant, Julien Renaud, Romain Delunel, David C. Tank, and Sébastien Lavergne.
815 2017. "Riders in the Sky (Islands): Using a Mega-Phylogenetic Approach to Understand Plant Species
816 Distribution and Coexistence at the Altitudinal Limits of Angiosperm Plant Life." *Journal of Biogeography*
817 44 (11). Wiley: 2618–30.
- 818 McPeck, Mark A. 2017. *Evolutionary Community Ecology*. Princeton University Press.
- 819 Meester, Luc De, Kristien I. Brans, Lynn Govaert, Caroline Souffreau, Shinjini Mukherjee, Hélène Vanvelk,
820 Konrad Korzeniowski, et al. 2018. "Analyzing Eco-Evolutionary Dynamics - the Challenging Complexity of
821 the Real World." *Functional Ecology*. Wiley.
- 822 Méléard, S., and S. Roelly. 1993. "Interacting Measure Branching Processes. Some Bounds for the Support." *Stochastics and Stochastic Reports* 44 (1-2). Informa UK Limited: 103–21.

- 824 Mubayi, Anuj, Christopher Kribs, Viswanathan Arunachalam, and Carlos Castillo-Chavez. 2019. "Studying
825 Complexity and Risk Through Stochastic Population Dynamics: Persistence, Resonance, and Extinction in
826 Ecosystems." In *Handbook of Statistics*, 157–93. Elsevier.
- 827 Nowak, Martin A. 2006. *Evolutionary Dynamics: Exploring the Equations of Life*. Belknap Press.
- 828 Nuismer, Scott L., Michael Doebeli, and Danny Browning. 2005. "The Coevolutionary Dynamics of An-
829 tagonistic Interactions Mediated by Quantitative Traits with Evolving Variances." *Evolution* 59 (10). The
830 Society for the Study of Evolution: 2073.
- 831 Nuismer, Scott L., and Luke J. Harmon. 2014. "Predicting Rates of Interspecific Interaction from Phyloge-
832 netic Trees." Edited by Jerome Chave. *Ecology Letters* 18 (1). Wiley: 17–27.
- 833 Nuismer, Scott L., Pedro Jordano, and Jordi Bascompte. 2012. "Coevolution and the Architecture of
834 Mutualistic Networks." *Evolution* 67 (2). Wiley: 338–54.
- 835 Nuismer, Scott L., Benjamin J. Ridenhour, and Benjamin P. Oswald. 2007. "Antagonistic Coevolution
836 Mediated by Phenotypic Differences Between Quantitative Traits." *Evolution* 61 (8). Wiley: 1823–34.
- 837 Nuismer, Scott L., Bob Week, and Marcelo A. Aizen. 2018. "Coevolution Slows the Disassembly of Mutu-
838 alistic Networks." *The American Naturalist* 192 (4). University of Chicago Press: 490–502.
- 839 Nuland, Michael E. Van, Ian M. Ware, Joseph K. Bailey, and Jennifer A. Schweitzer. 2019. "Ecosystem
840 Feedbacks Contribute to Geographic Variation in Plant-Soil Eco-Evolutionary Dynamics Across a Fertility
841 Gradient." Edited by Franziska Brunner. *Functional Ecology* 33 (1). Wiley: 95–106.
- 842 Page, Karen M., and Martin A. Nowak. 2002. "Unifying Evolutionary Dynamics." *Journal of Theoretical*
843 *Biology* 219 (1). Elsevier BV: 93–98.
- 844 Parachnowitsch, Amy L, Jessamyn S Manson, and Nina Sletvold. 2018. "Evolutionary Ecology of Nectar."
845 *Annals of Botany* 123 (2). Oxford University Press (OUP): 247–61.
- 846 Patel, Swati, and Reinhard Bürger. 2019. "Eco-Evolutionary Feedbacks Between Prey Densities and Linkage
847 Disequilibrium in the Predator Maintain Diversity." *Evolution* 73 (8). Wiley: 1533–48.
- 848 Perkins, Edwin A. 1991. "Conditional Dawson-Watanabe Processes and Fleming-Viot Processes." In *Semi-
849 nar on Stochastic Processes, 1991*, 143–56. Birkhäuser Boston.
- 850 ———. 1992. "Measure-Valued Branching Diffusions with Spatial Interactions." *Probability Theory and*
851 *Related Fields* 94 (2). Springer Science; Business Media LLC: 189–245.
- 852 ———. 1995. *On the Martingale Problem for Interactive Measure-Valued Branching Diffusions*. Amer
853 Mathematical Society.
- 854 Price, George R. 1970. "Selection and Covariance." *Nature* 227 (5257). Springer Nature: 520–21.
- 855 Queller, David C. 2017. "Fundamental Theorems of Evolution." *The American Naturalist* 189 (4). University
856 of Chicago Press: 345–53.
- 857 Reimers, Mark. 1989. "One Dimensional Stochastic Partial Differential Equations and the Branching Mea-
858 sure Diffusion." *Probability Theory and Related Fields* 81 (3). Springer Nature: 319–40.
- 859 Robertson, Alan. 1966. "A Mathematical Model of the Culling Process in Dairy Cattle." *Animal Science* 8
860 (1). Cambridge University Press: 95–108.
- 861 Roughgarden, Joan. 1979. *Theory of Population Genetics and Evolutionary Ecology: An Introduction*.
862 Macmillan.
- 863 Rudman, Seth M., Matthew A. Barbour, Katalin Csilléry, Phillip Gienapp, Frederic Guillaume, Nelson G.
864 Hairston Jr, Andrew P. Hendry, et al. 2017. "What Genomic Data Can Reveal About Eco-Evolutionary
865 Dynamics." *Nature Ecology & Evolution* 2 (1). Springer Nature: 9–15.
- 866 Schuster, Peter, and Karl Sigmund. 1983. "Replicator Dynamics." *Journal of Theoretical Biology* 100 (3).
867 Elsevier BV: 533–38.

- 868 Skovmand, Lotte H., Charles C.Y. Xu, Maria R. Servedio, Patrik Nosil, Rowan D.H. Barrett, and Andrew
869 P. Hendry. 2018. “Keystone Genes.” *Trends in Ecology & Evolution* 33 (9). Elsevier BV: 689–700.
- 870 Sterner, R.W., and J.J. Elser. 2008. “Ecological Stoichiometry: Overview.” In *Encyclopedia of Ecology*,
871 1101–16. Elsevier.
- 872 Taylor, Peter D., and Leo B. Jonker. 1978. “Evolutionary Stable Strategies and Game Dynamics.” *Mathe-*
873 *matical Biosciences* 40 (1-2). Elsevier BV: 145–56.
- 874 Tilman, David. 1982. *Resource Competition and Community Structure*. Princeton University Press.
- 875 Turelli, Michael. 1984. “Heritable Genetic Variation via Mutation-Selection Balance: Lerchs Zeta Meets the
876 Abdominal Bristle.” *Theoretical Population Biology* 25 (2). Elsevier: 138–93.
- 877 ———. 1986. “Gaussian Versus Non-Gaussian Genetic Analyses of Polygenic Mutation-Selection Balance.”
878 In *Evolutionary Processes and Theory*, 607–28. Academic Press.
- 879 ———. 2017. “Commentary: Fisher’s Infinitesimal Model: A Story for the Ages.” *Theoretical Population*
880 *Biology* 118 (December). Elsevier BV: 46–49.
- 881 Turelli, Michael, and Nick Barton. 1994. “Genetic and statistical analyses of strong selection on polygenic
882 traits: what, me normal?” *Genetics* 138 (3): 913–41.
- 883 Walsh, John B. 1986. “An Introduction to Stochastic Partial Differential Equations.” In *Lecture Notes in*
884 *Mathematics*, 265–439. Springer Berlin Heidelberg.
- 885 Watanabe, Shinzo. 1968. “A Limit Theorem of Branching Processes and Continuous State Branching
886 Processes.” *Journal of Mathematics of Kyoto University* 8 (1). Duke University Press: 141–67.
- 887 Wright, Sewall. 1931. “Evolution in Mendelian Populations.” *Genetics* 16 (2). Genetics: 97–159.
- 888 Xiao, Xiao, Daniel J. McGlinn, and Ethan P. White. 2015. “A Strong Test of the Maximum Entropy Theory
889 of Ecology.” *The American Naturalist* 185 (3). University of Chicago Press: E70–E80.